A MUTATIONAL ANALYSIS OF MOTIFS IN *Eco* KI COMMON TO ADENINE METHYLTRANSFERASES

Damion F. Willcock

A thesis presented for the degree of Ph.D

Institute of Cell and Molecular Biology
University of Edinburgh

September 1994

ABSTRACT

Type I restriction-modification systems are complex multifunctional enzymes comprising three polypeptides; HsdR, M and S. All three polypeptides are required to form a restriction enzyme but M and S alone are sufficient to form an N6 adenine methyltransferase. The HsdM polypeptide of the type I system *Eco* KI contains motifs characteristic of N6 adenine methyltransferases (N/D P P F/Y) and of methyltransferases in general (F X G X G). These motifs may identify amino acid sequences critical to methyltransferase function.

This work describes an *in vivo* and *in vitro* analysis of site-directed mutations within these two motifs and is the first report of such changes within a type I system.

Biochemical analysis identifies the N/D P P F/Y region as critical to catalysis and in close proximity to the S-adenosylmethionine-binding site. Mutations which remain within the overall consensus do not necessarily retain activity and hence indicate a degree of stringency associated with this motif.

A mutation within the F X G X G motif abolishes S-adenosylmethionine binding but retains DNA binding activity. In addition this mutant is temperature sensitive.

ABBREVIATIONS

AdoMet	S-adenosylmethionine
AMPS	ammonium persulphate
ANS	(1, 8-anilino-napthalene sulphonic acid)
bp	base pair(s)
BSA	bovine serum albumin
CE	conserved element
cpm	counts per minute
dNTP	deoxynucleoside triphosphate
ddNTP	dideoxynucleoside triphosphate
DTT	dithiothreitol
EDTA	diaminoethanetetraacetic acid
e.o.p.	efficiency of plating
g	relative centrifugal force
H-bonds	hydrogen bonds
hsd	host specificity determinant (host specific defence)
IPTG	isopropyl-β-D-thiogalactoside
kb	kilobase(s)
m.o.i	multiplicity of infection
O.D.	optical density
PEG	polyethylene glycol
pfu	plaque forming unit
RF	replicative form
SDS	sodium dodecyl sulphate
SSC	standard saline citrate
TEMED	N, N, N', N'-teramethyl ethylene diamine
TRD	target recognising domain
ts	temperature sensitive
Δ	deletion

AMINO ACIDS

Name	Three Letter Code	Single Letter Code
alanine	Ala	A
arginine	Arg	R
asparagine	Asn	N
aspartate	Asp	D
cysteine	Cys	C
glutamate	Glu	E
glutamine	Gln	Q
glycine	Gly	G
histidine	His	H
isoleucine	Ile	I
leucine	Leu	L
lysine	Lys	K
methionine	Met	M
phenylalanine	Phe	F
proline	Pro	P
serine	Ser	S
threonine	Thr	T
tryptophan	Trp	W
tyrosine	Tyr	Y
valine	Val	V

CONTENTS

Chapter One

INTRODUCTION

1.1 Introduction

Restriction and modification systems (R-M) are pairs of opposing enzyme activities found in bacteria. Their primary function is probably to protect the host from foreign DNA, particularly that of bacteriophage. However an alternative role may be to promote recombination as a means of generating variability and hence increase the rate of evolution (Price and Bickle, 1986).

R-M systems were first recognised in the early 1950's and their discovery led to one of the most important breakthroughs in molecular biology. One of the first systems to be described affected the growth of bacteriophage (phage) λ grown on different bacterial strains (Bertani and Weigle, 1952). When grown on *E.coli* K-12, λ has an equal efficiency of plating (e.o.p.) on *E.coli* K-12 and *E.coli* C. However when grown on *E.coli* C, λ plates with a much lower efficiency on *E.coli* K-12 than on *E.coli* C. Non-adsorption of phage was excluded as a possible reason for this restriction of growth. It was noted that the effect was a non-heritable alteration stemming directly from passage through a new host and not from mutation. Each host altered the phage in a characteristic way, independent of the previous host history of the phage.

The restriction and modification of phage was redefined in molecular terms when it was demonstrated that the restriction of phage growth imposed by the host was associated with the breakdown of phage DNA (Lederberg, 1957). Using ^{32}P labelled DNA, it was found that when a phage infected a restricting host, the label was soon located within the cell but thereafter was increasingly found in the medium. This indicated degradation of the DNA and diffusion of the products out of the cell. Furthermore, single rounds of infection using isotopically labelled modified DNA which was replicated in a non-modifying host, demonstrated that the modification was maintained only in those progeny with one or both strands of parental DNA (Arber and Dussoix, 1962). As well as certain phage DNAs, bacterial DNA transferred between cells by conjugation was also found to be subjected to R-M systems (Arber and Morse, 1965).

The genetic distinction between restriction and modification was demonstrated by the isolation of mutants which no longer restricted an unmodified P1 phage (Glover *et al.*, 1963). This phenotype was designated restriction$^-$ (r^-).

Some of the mutants retained modification activity (r^-m^+), whereas others had lost both functions (r^-m^-). Similarly in the K12 and B systems, r^- mutants were isolated, some of which were r^-m^+ and others which were r^-m^- (Wood, 1966).

A further experiment indicated that methionine (or a derivative) was involved in the modification process (Arber, 1965). Infection of a methionine auxotrophic *E.coli* K-12 host with λ, in the absence of methionine, generated progeny of which the majority were unmodified. No other amino acid tested in this way gave the same lack of modification. Experiments with phage fd using ^{14}C methylmethionine as a label demonstrated a correlation between B-specific modification and the N6 methylated form of adenine, 6-methylaminopurine (Arber, 1964). The level of incorporation into this form was very low, corresponding to perhaps one or two residues per fd genome. These experiments indicated that methionine serves as a methyl donor in the modification process, presumably via AdoMet.

Following the purification of several restriction endonucleases it became clear that the enzymes could be subdivided, based on their cofactor requirements (Takano *et al.*, 1966; Meselson and Yuan, 1968; Linn and Arber, 1968; Smith and Wilcox, 1970). Initially the enzymes were split into two groups, types I (ATP dependent) and II (Boyer, 1971). This original classification was revised when it was shown that the ATP dependent enzymes could be subdivided into those which hydrolyse ATP (type I) and those which do not (type III) (Haberman, 1974; Kauc and Piekarowicz, 1978).

1.2 Restriction and Modification Systems

1.2.1 Type II Restriction-Modification Systems

Of the three systems the type II R-M systems are the most widespread and simple (reviewed in Wilson and Murray, 1991; Bickle and Krüger, 1993). Mg^{2+} is required for restriction and AdoMet is required for modification. The restriction endonucleases act as homodimers, catalysing the double stranded endonucleolytic cleavage of DNA at a fixed position within a specific symmetrical target sequence. Target sites are commonly 6bp in length. A subset, the type IIs systems (*s* hifted cleavage), have an asymmetric target sequence and cleave outside of the recognition

sequence at a defined distance of up to 20bp. The methyltransferases of type II systems act as monomers and modify symmetrically located nucleotides within the target sequence. The methylated nucleotide may be adenine, giving N6-methyladenine (m6A), or cytosine, giving either C5-methylcytosine (m5C) or occasionally N4-methylcytosine (m4C). Methylation of the correct nucleotide on one or both strands of the recognition sequence renders the sequence resistant to cleavage.

1.2.2 Type III Restriction-Modification Systems

Type III systems comprise only four members and have an intermediate level of complexity (reviewed in Wilson and Murray, 1991; Bickle and Krüger, 1993). These systems comprise two subunits, Mod and Res. The Mod subunit alone acts as a methyltransferase and together with the Res subunit forms both a methyltransferase and endonuclease. Mg^{2+} is required for both activities. AdoMet is required for methylation and also stimulates restriction. ATP is required for restriction but is not hydrolysed. The asymmetric target sequences of type III systems are 5-6 bp and cleavage occurs to one side, approximately 25bp away. Methylation (N6) occurs on only one strand with the resulting paradox that totally unmodified DNA will arise following replication of this hemimethylated DNA. This anomaly is explained by the fact that restriction requires the presence of two unmethylated sites in opposite orientations. The sites of the host DNA all lie in the same orientation and are therefore refractory to cleavage (Meisel *et al.*, 1992).

1.2.3 Type I Restriction-Modification Systems

Type I systems are the most complex of the three types. In addition to restriction (R) and modification (M) subunits, they possess a third subunit, S, responsible for the DNA specificity of both restriction and modification. While all three gene products are required for restriction, only the M and S subunits are required for modification (Lautenberger and Linn, 1972; Sain and Murray, 1980; Suri *et al.*, 1984a).

Type I systems methylate the N6 position of adenine in asymmetric bipartite recognition sequences. The two parts of the recognition sequence are separated by a spacer of non-specific sequence but fixed length. The two target adenines (**A**) are in opposite strands and their spacing is such that they are separated by approximately one helical turn. The recognition sequence for the *Eco* KI system for example is:

5'-A**A**C(N_6)GTGC
TTG(N_6)C**A**CG-5' (Kan *et al.*, 1979)

See table 1.1 for a list of the type I systems (and their target recognition sites) identified in enteric bacteria to date.

Table 1.1

Type I Restriction-Modification Systems identified in Enteric Bacteria

Family	Enzyme	Target	Reference
IA	***Eco* KI**	**AAC (N_6) GTGC**	**Kan *et al.*, 1979**
	Eco BI	TGA (N_6) TGCT	Lautenberger *et al.*, 1978
			Ravetch *et al.*, 1978
			Sommer & Schaller, 1979
	Eco DI	AAC (N_6) GTGY	Nagaraja *et al.*, 1985a
	Sty LTIII	GAG (N_6) RTAYG	Nagaraja *et al.*, 1985b
	Sty SPI	AAC (N_6) GTRC	Nagaraja *et al.*, 1985b
	Eco R5I	unknown*	V.A.Barcus, unpublished
	Eco R23I	unknown*	V.A.Barcus, unpublished
IB	*Eco* AI	AAC (N_7) GTCA	Suri *et al.*, 1984b
	Eco EI	AAC (N_6) ATGC	Cowan *et al.*, 1989
	Eco R42I	unknown*	V.A.Barcus, unpublished
	C fr AI	GCA (N_8) GTGG	Kannan *et al.*, 1989
	Sty STI	unknown*	A.J.B.Campbell &
			N.E.Murray, unpublished
	Sty SKI	unknown*	D.Ternent & N.E.Murray,
			unpublished
IC	*Eco* R124I	GAA (N_6) RTCG	Price *et al.*, 1987
	Eco DXXI	TCA (N_7) RTTC	Gubler *et al.*, 1992
	Eco prrI	CCT (N_7) RTGC	Tyndall *et al.*, 1994
ID?	*Sty* SBLI	unknown*	A.J.B.Campbell &
			N.E.Murray, unpublished
	Eco R9I	unknown*	A.J.B.Campbell &
			N.E.Murray, unpublished

* Target sequences unknown but each enzyme shown to differ from other members of the same family.

1.3 The Characteristics of Type I Restriction-Modification Systems

1.3.1 The Genetics

Complementation studies in the late 1960's using F' plasmids to generate merodiploids indicated that type I systems are encoded by 3 genes; *hsdR* , *M* and *S* (Boyer and Roulland-Dussoix, 1969; Glover, 1970). In addition to demonstrating the involvement of three genes in type I systems, these studies gave the first indication that different systems might be closely related to each other. Complementation tests between mutant derivatives of the genes encoding the systems *Eco* KI and *Eco* BI demonstrated that the enzyme subunits of these systems are interchangeable (Boyer and Roulland-Dussoix, 1969).

The *hsd* genes of the *Eco* KI system (*hsdK*) were first cloned, in phage λ, using a combination of *in vivo* and *in vitro* techniques (Sain and Murray, 1980). Analysis of λ*hsdK* phages identified three genes, *hsdR* , *M* and *S* , encoding polypeptides with predicted molecular weights of approximately 130000, 62-65000 and 50000 respectively (Sain and Murray, 1980). Analysis of deletion phages identified one transcriptional unit for the expression of both *hsdM* and *hsdS* , and another for *hsdR* . On the basis of transcription of the *hsdK* region with respect to λ promoters, the orientation of the transcriptional units was identified as being in the same direction, with the first unit transcribing *hsdR* and the second unit transcribing *hsdM* and *S* .

Sequencing of the *hsdK* region identified two promoters, confirming the genetic data of Sain and Murray (Loenen *et al.*, 1987). *hsdR* and *hsdM* are separated by almost 0.5kb whereas the termination codon of *hsdM* overlaps the initiation codon of *hsdS* . This organisation may be of physiological importance since it could permit independent transcription of the genes necessary for modification but there is no evidence of regulation at this level (Loenen *et al.*, 1987). However, in a study of *hsdK* expression following conjugal transfer, a delay in restriction has been detected. Following such transfer, endonuclease activity was not detected until after approximately fifteen generations (Prakash-Cheng and Ryu, 1993). Furthermore, the product of a fourth gene, *hsdC* , has been demonstrated to be necessary in enabling the conjugal transfer of functional *hsdK* genes into a cell, the DNA of which is not protected by modification. A mutation in *hsdR* (giving an

r^- phenotype) removes the barrier present in an *hsdC* $^-$ recipient (Prakash-Cheng *et al.*, 1993). Similar delays in expression have been observed in type III systems (Arber and Dussoix, 1962) and genes controlling restriction activity have been associated with several type II systems (Tao *et al.*, 1991; Ives *et al.*, 1992).

Type I R-M systems can be grouped into three families (a fourth family, ID, has recently been discovered, Campbell, A.J.B., and N.E. Murray, unpublished). This aspect makes them particularly attractive with respect to both the investigation of bacterial evolution and the analysis of molecular interactions. Preliminary DNA hybridisation studies supported early complementation data by demonstrating extensive hybridisation of the *Eco* KI *hsdR* and *M* genes to the DNA of *E.coli* B (Sain and Murray, 1980). Furthermore, no hybridisation was detected between the *Eco* KI probe and the DNA of *E.coli* C, in agreement with genetic tests which indicated the absence of an R-M system in the latter strain. More extensive DNA hybridisation and immunological studies were used to examine the relatedness of several type I systems (Murray *et al.*, 1982). Strong hybridisation was detected between an *Eco* KI probe and the DNA of *E.coli* B and weaker but nonetheless distinct hybridisation between the *Eco* KI probe and the DNA of the *Salmonella* strains *Sty* LTIII, and *Sty* SPI. This indicated that these systems are part of one family, the K or IA family. However no hybridisation was detected between the *Eco* KI probe and the DNA of *E.coli* A despite the fact that the *Eco* AI genes are allelic to those of *Eco* KI and that the *Eco* AI system displays typical type I characteristics. This indicated that *Eco* AI represents a different family, the A or IB family.

In a further screening experiment based on DNA hybridisation a wider spectrum of enteric bacteria were examined for possession of *hsd* genes (Daniel *et al.*, 1988). Using IA and IB related probes no hybridisation was detected in the majority of *E.coli* isolates but in *Citrobacter freundii* hybridisation to the IB specific probe was demonstrated. Cloning of this region of the *C.freundii* genome provided a functional IB related restriction system of novel specificity. Similarly, using the same technique, other members of the IB family have been discovered, for example, the *Eco* EI system and, more recently, the *Eco* R42I system (Fuller-Pace *et al.*, 1985; Barcus,V.A. unpublished).

A third type I family, the R124 or IC family, is mainly plasmid encoded (Price *et al.*, 1987). To date this family contains three members; *Eco* R124I, *Eco* DXXI and *Eco* prrI (Price *et al.*, 1987; Gubler *et al.*, 1992; Tyndall *et al.*, 1994).

Between families there is very little similarity in nucleotide sequence. Likewise, between families there is very little amino acid sequence similarity apart from motifs common to ATP dependent helicases (see section 1.3.2) and to DNA methylases (see section 1.4)(Murray *et al.*, 1993). Within the IA family, *hsdR* and *hsdM* are largely conserved throughout their lengths at both the nucleotide and amino acid sequence levels. The M polypeptides of *Eco* KI and *Eco* BI for example have 96.6% sequence identity. Within the IB family the level of similarity is less, the M polypeptides of *Eco* AI and *Eco* EI having 89.8% identity. Interestingly, analysis of silent substitutions between the *hsdM* genes of *Eco* AI and *Eco* EI indicates a high level of divergence. The level is even higher than that observed between *E.coli* and *Salmonella* specified IA-family *hsdM* genes. Murray *et al.*, 1993, have suggested that this finding is consistent with a comparison between species rather than within species of bacteria.

Type I *hsdS* genes have a characteristic structure comprising two large regions of high sequence divergence (figure 1.1). These 'variable' regions alternate with sequences of much greater similarity, the conserved regions (Gough and Murray, 1983). It was suggested that the two variable regions might be the target recognition domains (TRDs), responsible for determining the specificity of the two halves of the bipartite recognition sequence (Gough and Murray, 1983). Systems which share common elements in their recognition sites have variable regions with high sequence identity (Fuller-Pace and Murray, 1986). The trinucleotide component 5'-AAC, for example, is common to both *Eco* KI and *Sty* SPI and their amino variable regions share over 90% identity. Furthermore, Bullas *et al.*, 1976, isolated a novel specificity which was presumed to have resulted from a cross-over event between two IA *S* genes from the systems *Sty* SPI and *Sty* LTIII. The recombination event was localised between the two large variable regions and was found to generate a novel *S* gene comprising the 5' half of *Sty* SPI *hsdS* and the 3' half of *Sty* LTIII *hsdS*. (Fuller-Pace *et al.*, 1984; Fuller-Pace and Murray, 1986). The specificity of the resulting system, *Sty* SQ, was found to comprise the trimeric component of *Sty* SPI and the tetrameric component of *Sty* SBI (Nagaraja *et al.*, 1985c). The reciprocal recombinant *Sty* SJI was constructed and as expected the

resulting *hsdS* gene encoded a polypeptide which specified the trimeric component from *Sty* SBI and the tetrameric component from *Sty* SPI (Gann *et al.*, 1987). Similar experiments have been performed with the type IC systems *Eco* R124I and *Eco* DXXI to generate recombinants with novel combinations of TRDs (figure 1.2).

In type IA *hsdS* genes the TRDs are separated by a region which is conserved within the IA family, the central conserved region. An additional conserved region is found at the carboxy terminii of the IA *hsdS* genes. The IB and IC families, in addition to having central and carboxy conserved regions, also have a conserved region at the amino terminus (figure 1.1). The IA and IB *hsdS* genes contain repeated sequences largely within their conserved regions in a manner which indicates that the genes may have arisen as a result of gene duplication (Gough and Murray, 1983; Kannan *et al.*, 1989). An ancestral gene with a single recognition domain could have duplicated to generate the current structural plan. Subsequent divergence of the two DNA recognition domains may have left the repeats as the only visible remnants of the duplication event.

Analysis of the *hsdS* genes of the type IC systems *Eco* R124I and the closely related *Eco* R124II indicates one role for the central conserved region as determining the length of the non-specific spacer in the target recognition sequence. The two systems differ only in the length of this spacer. *Eco* R124I specifies a 6 bp spacer whereas *Eco* R124II specifies a 7 bp spacer (Price *et al.*, 1987). Sequencing of the two systems reveals that *Eco* R124II differs in having three copies of a 12 bp repeat located between the two large variable regions of *hsdS*. *Eco* R124I has two copies of this repeat. The region of the polypeptide containing these repeats serves as a linking element between the two DNA-binding domains (figure 1.2). The extra four amino acids resulting from the third copy of the repeat is sufficient to move the domains an extra base pair apart resulting in a new specificity. Site-directed mutagenesis confirmed that the exact sequence of the repeat was not important for activity but that its length was critical in providing productive interactions (Gubler and Bickle, 1991).

The similarity of the *hsdS* conserved regions within families has prompted the suggestion that they may be involved in subunit-subunit interaction (Gann *et al.*, 1987; Kannan *et al.*, 1989). This idea is supported by the analysis of novel type IC systems which have been generated by inactivation of the 3' regions of *Eco* DXXI

hsdS and *Eco* R124I *hsdS* (Meister *et al.*, 1993; Abadjieva *et al.*, 1993). The resulting *S* genes retain a TRD for only the trinucleotide component of their respective target sequences in addition to retaining the central and N-terminal conserved regions. Interestingly the truncated genes result in functional systems. The target sequences of these systems have the usual bipartite structure but the trinucleotide component of the target site provides both of the recognition elements by flanking the spacer in a symmetrical arrangement (figure 1.2). The mutant methylases would seem to comprise two of the deleted HsdS polypeptides (stoichiometry M_2S_2) rather than one full length HsdS as in a wild-type type I methylase (stoichiometry M_2S).

Figure 1.1

The Conserved and Variable Regions of Type I HsdS Polypeptides.

Figure 1.1

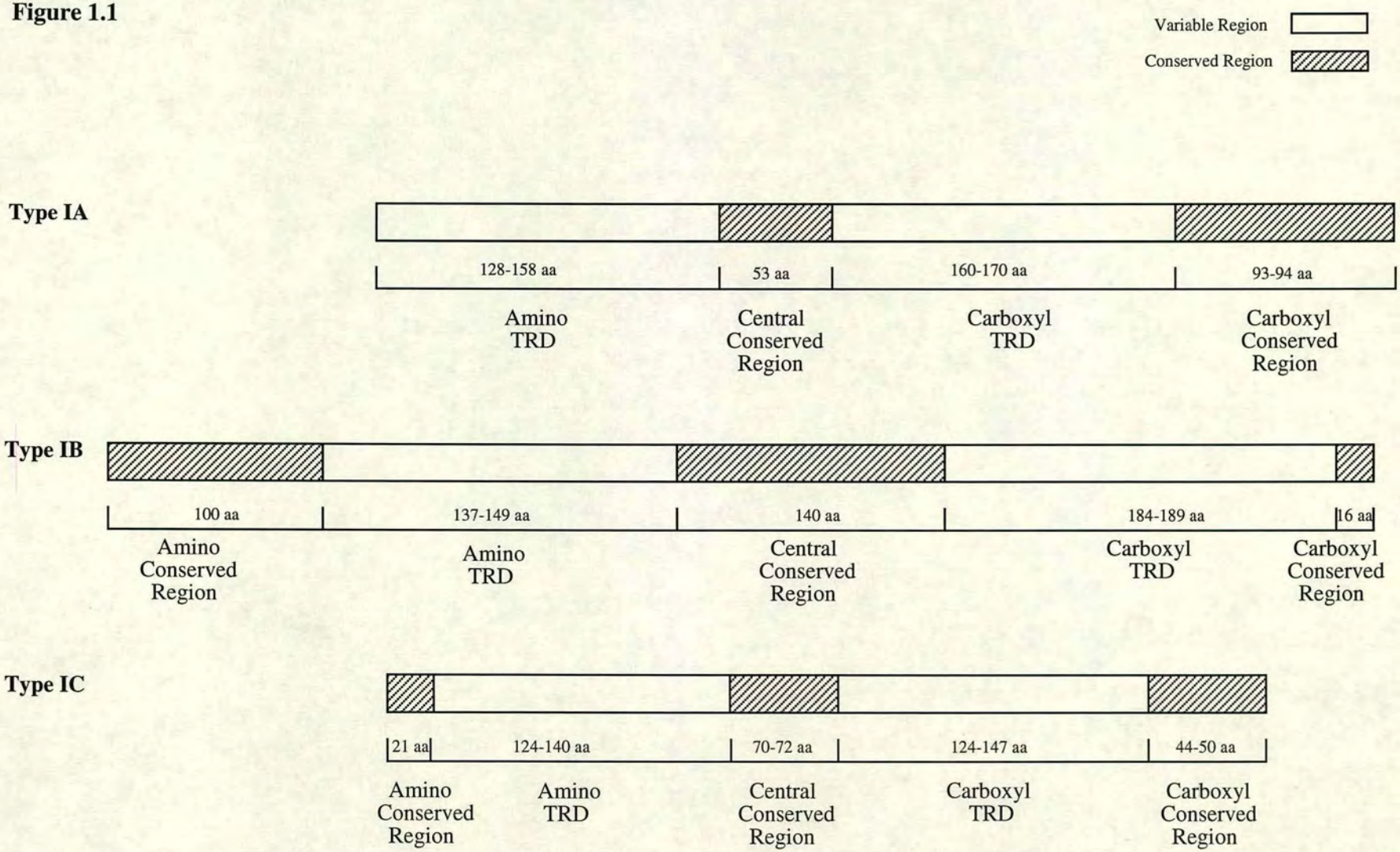

Variable Region
Conserved Region
Type IA
128-158 aa
53 aa
160-170 aa
93-94 aa
Amino TRD
Central Conserved Region
Carboxyl TRD
Carboxyl Conserved Region
Type IB
100 aa
137-149 aa
140 aa
184-189 aa
16 aa
Amino Conserved Region
Amino TRD
Central Conserved Region
Carboxyl TRD
Carboxyl Conserved Region
Type IC
21 aa
124-140 aa
70-72 aa
124-147 aa
44-50 aa
Amino Conserved Region
Amino TRD
Central Conserved Region
Carboxyl TRD
Carboxyl Conserved Region

Figure 1.2

Type IA and IB HsdS Polypeptides and the Target Recognition Sequences which they specify. Those marked (#) are natural isolates. The remainder are engineered constructs.

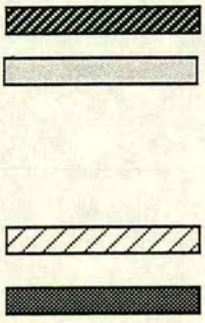
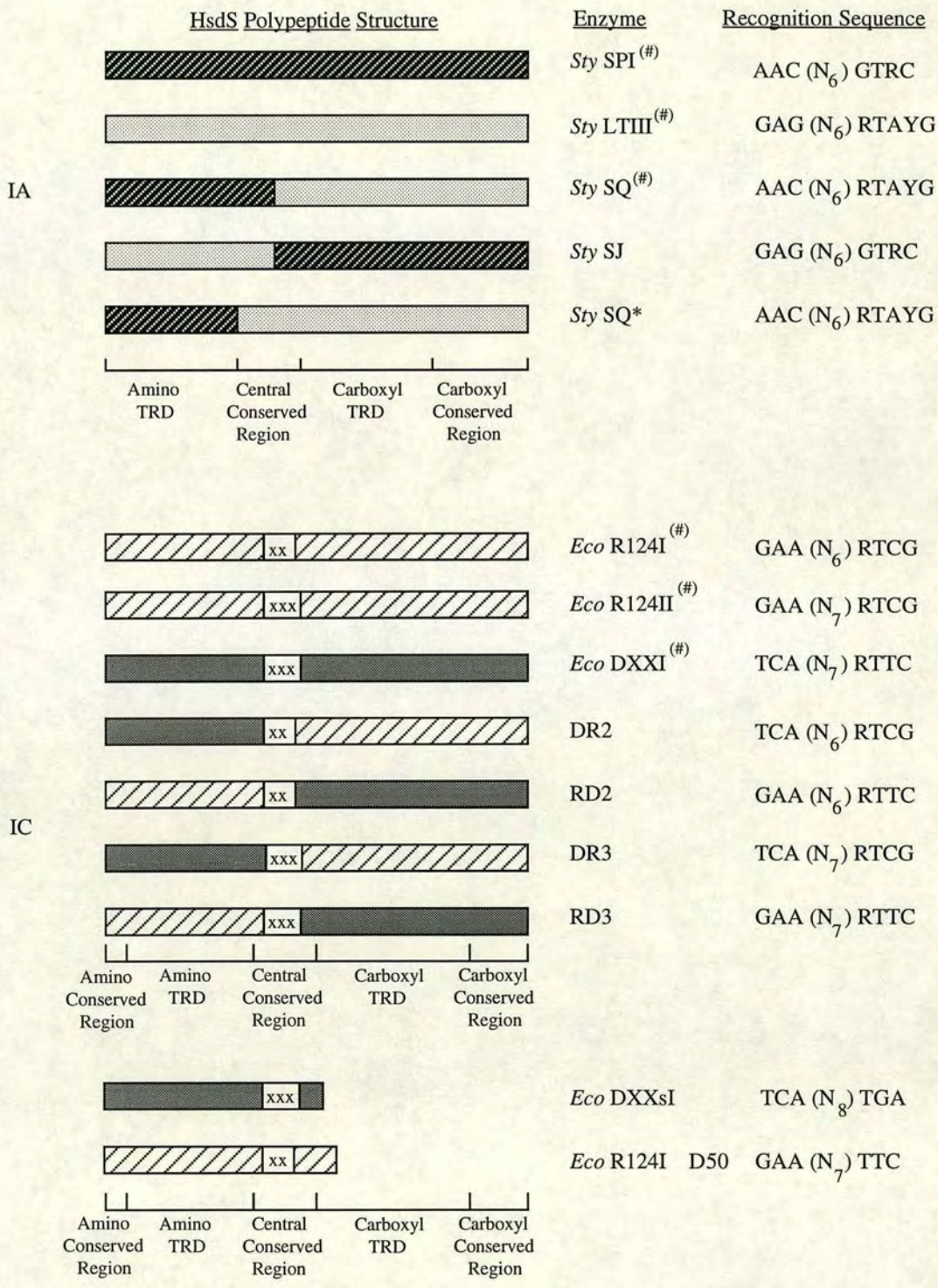


Figure 1.2


HsdS Polypeptide Structure
Enzyme
Recognition Sequence
IA
Sty SPI(#)
AAC (N6) GTRC
Sty LTIII(#)
GAG (N6) RTAYG
Sty SQ(#)
AAC (N6) RTAYG
Sty SJ
GAG (N6) GTRC
Sty SQ*
AAC (N6) RTAYG
Amino TRD
Central Conserved Region
Carboxyl TRD
Carboxyl Conserved Region
IC
xx
xxx
Eco R124I(#)
GAA (N6) RTCG
Eco R124II(#)
GAA (N7) RTCG
Eco DXXI(#)
TCA (N7) RTTC
DR2
TCA (N6) RTCG
RD2
GAA (N6) RTTC
DR3
TCA (N7) RTCG
RD3
GAA (N7) RTTC
Amino Conserved Region
Amino TRD
Central Conserved Region
Carboxyl TRD
Carboxyl Conserved Region
Eco DXXsI
TCA (N8) TGA
Eco R124I D50
GAA (N7) TTC

1.3.2 The Reaction Mechanism

Much of the information on type I systems comes from the IA enzymes *Eco* KI and *Eco* BI. The systems function as both endonucleases and methyltransferases. In addition they possess ATPase activity.

The first step in the reaction mechanism (figure 1.3) is the binding of AdoMet. This leads to an activated form of the enzyme, *Eco* K*, which is able to bind DNA firstly in a non-specific manner and then at the specific recognition site (Yuan *et al.*, 1975). At methylated sites ATP stimulates the dissociation of the enzyme from the DNA (Bickle *et al.*, 1978). At hemimethylated sites ATP stimulates methylation of the unmethylated strand. In the case of unmethylated sites ATP converts the enzyme into a new form, *Eco* K^+, which is characteristic in the following respects; it does not bind AdoMet, it forms a filter-binding complex with DNA and it shows an apparent change in conformation (Bickle *et al.*, 1978). DNA is then translocated past the enzyme with concurrent ATP hydrolysis (Yuan *et al.*, 1980). DNA cleavage occurs at a variable position of up to several kb from the target recognition site and ATP hydrolysis continues for some time afterwards.

Several hypotheses have been proposed to explain the mechanism by which the restriction enzyme cleaves DNA at positions distant to the recognition site. One suggestion is that the enzyme dissociates from the recognition site and then interacts with a random site on either the same or a different DNA molecule. This model is not likely since the enzyme is unable to cleave modified DNA in *trans* when incubated with a mixture of modified and unmodified DNA (Meselson and Yuan, 1968). Another model is one in which the enzyme tracks along the DNA away from its target site until it encounters a potential cleavage site. However this is also unlikely since the enzyme appears to remain bound at the recognition site even after cleavage (Bickle *et al.*, 1978).

Studies of the restriction mechanism of *Eco* KI and *Eco* BI have indicated the formation of looped DNA structures (Yuan *et al.*, 1980; Endlich and Linn, 1985). A model in which the enzyme has two DNA-binding sites has been suggested which accounts for the formation of such structures (Yuan *et al.*, 1980). The first site is the specific recognition sequence, to which the enzyme binds in a particular orientation. The second site is not sequence specific and only becomes available, perhaps by

random collision, once ATP has induced the conformational change to *Eco* K^+. The binding of the second DNA site might initiate DNA translocation past the enzyme, generating supercoiled loops.

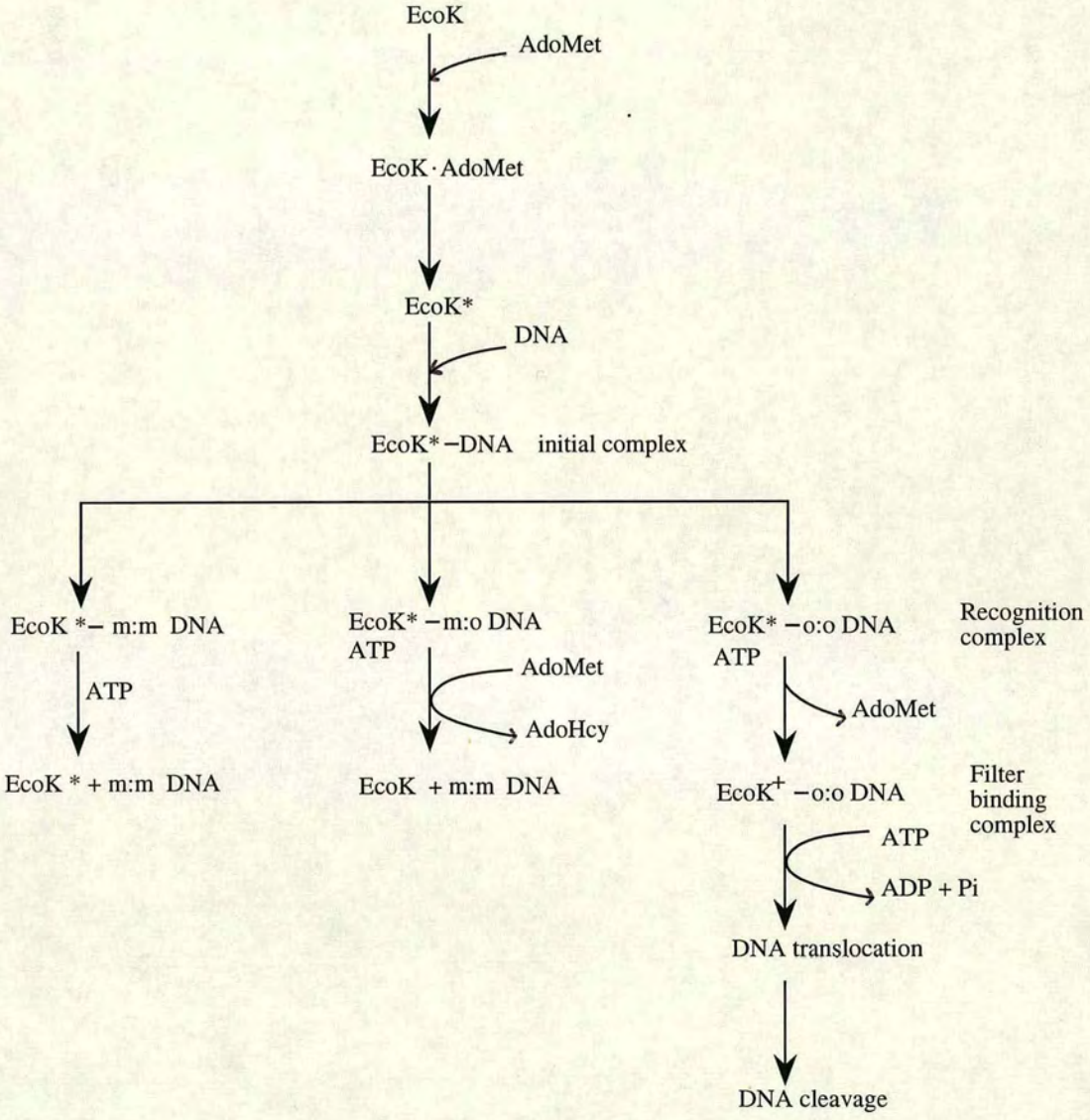
A further development of this model attempts to explain the 'apparently' random selection of a cleavage site and proposes that cleavage is dependent on the interaction between two enzyme molecules (Studier and Bandyopadhyay, 1988). This model accounts for the orderly progression of DNA fragments observed when phage T7 DNA is digested with the *Eco* KI restriction enzyme. The positions of the cleavage sites cluster at the midpoints between adjacent recognition sites. This can be explained if an enzyme bound at the recognition site translocates DNA toward itself from both directions and when translocation causes neighbouring enzymes to meet, DNA cleavage occurs. At relatively low enzyme to DNA ratios, DNA containing only one site is not cleaved. However, at relatively high enzyme to DNA ratios, DNA can be cleaved by the interaction between the specifically bound enzyme and excess unbound enzyme. Murray *et al.*, 1973, have shown that a linear substrate with only one site is only restricted in the presence of excess enzyme. However observations of the *Eco* KI restriction of circular DNA molecules containing only one site may suggest that restriction of such molecules can be achieved by the specifically bound enzyme alone, with no dependency on a second enzyme molecule (Horiuchi and Zinder, 1972; Weiserova *et al.*, 1993). This would suggest that linear molecules and circular molecules with only one site may be cleaved at rates of different efficiency.

Type I R subunits contain several motifs characteristic for both ATP-binding sites (Walker *et al.*, 1982) and ATP-dependent helicases, in particular the so called 'DEAD' box (Gorbalenya and Koonin, 1991). This is consistent with a model in which ATP-dependent translocation of DNA occurs as part of the restriction activity.

Figure 1.3

The Reaction Mechanism of the Restriction Endonuclease *Eco* KI. The action of the enzyme is dependent upon the methylation state of the target site: if the target site is fully methylated (m:m), the enzyme dissociates: if the target site is hemimethylated (o:m), the enzyme methylates the unmethylated strand: if the target site is unmethylated (o:o), restriction occurs (Burckhardt *et al.*, 1981).

Figure 1.3


EcoK
AdoMet
EcoK · AdoMet
EcoK*
DNA
EcoK* –DNA initial complex
EcoK *– m:m DNA
ATP
EcoK * + m:m DNA
EcoK* – m:o DNA
ATP
AdoMet
AdoHcy
EcoK + m:m DNA
EcoK* – o:o DNA
ATP
AdoMet
Recognition complex
EcoK+ – o:o DNA
Filter binding complex
ATP
ADP + Pi
DNA translocation
DNA cleavage

1.3.3 Type I R-M Systems: Methylase Activity

The methylation activity of type I systems is manifested not only by the restriction enzyme but also by a combination of M and S subunits alone (Lautenberger and Linn, 1972). In the *Eco* KI system the active form of the methylase is M_2S_1 although inactive forms of different stoichiometries (M_1S_1, M_1, M_2) are found upon purification (Dryden *et al.*, 1993).

The type IA methylases, as exemplified by the *Eco* KI and *Eco* BI enzymes, are the only prokaryotic methylases known to show a strong preference for a hemimethylated DNA target site (Suri and Bickle, 1985; Lautenberger and Linn, 1972) and in this respect they resemble the eukaryotic maintenance methylase (Bestor and Ingram, 1983). The rate of methylation of a hemimethylated DNA by the *Eco* KI methylase is 35x that for an unmodified DNA. This difference increases to 100x for methylation by the restriction enzyme (Suri and Bickle, 1985). The *Eco* R124I methylase, the best studied of the IC systems, has recently been shown to have a similar strong preference for hemimethylated DNA (Taylor *et al.*, 1993). The preference is less marked for methylation by the *Eco* R124I restriction enzyme (Price *et al*, 1987). Unlike the IA and IC systems, the IB enzymes have no preference for hemimethylated DNA (Suri and Bickle, 1985). *In vivo* the activity of the type IA methylases *Eco* KI and *Eco* BI can be enhanced by the λ gene product Ral which at the same time suppresses restriction (Zabeau *et al.*, 1980; Loenen and Murray, 1986). It has been suggested that confinement of Ral to action upon the IA family may be related to the stronger preferences for a hemimethylated DNA substrate (Kelleher *et al.*, 1991). In a IA system, if an infecting unmodified phage escapes restriction, it remains a poor substrate for KI modification since hemimethylated DNA is the much preferred substrate. Hence the phage remains sensitive to restriction and its probability of transmission is reduced (King, G., unpublished data). However, phage λ can overcome this barrier by producing Ral which effects KI specific modification of an unmethylated target. In IB systems Ral is redundant because unmethylated phage DNA is as good a substrate for modification as is hemimethylated DNA. *Eco* AI, the best studied of the IB systems, appears to have evolved a more efficient restriction system which perhaps compensates for this. An *Eco* AI target elicits approximately a five fold increase in the restriction of λ compared to an *Eco* KI target (Arber and Wauters-Willems, 1970; Murray *et al.*, 1973).

The isoelectric points for the S subunits of type I systems correlate with the three families. The S subunit isoelectric points of the type IA members are very basic (approximately pH 10), those of the IC family are less basic (pH 8.5) and those of the IB family are acidic (approximately pH 6). This pattern might be associated with the different preferences for the methylation state of the target site (Dryden *et al.*, 1993).

The effect of DNA methylation state and the effect of AdoMet on DNA binding by the *Eco* KI methylase has been investigated (Powell *et al.*, 1993). The methylase has two AdoMet-binding sites, one on each of the M subunits. The sites are equivalent and non-interacting, with dissociation constants of 2-3 μM. The complete restriction enzyme has previously been estimated to have at least 5 binding sites, all of which are involved in the methylation reaction and three of which are involved in restriction (Hadi *et al.*, 1975; Burckhardt *et al.*, 1981). However the R polypeptide, on its own, has been shown not to bind AdoMet (Dryden, D.T.F., personal communication). It is possible, though perhaps unlikely, that extra sites are created as a result of the interaction between R and the M and S polypeptides. The methylase has a higher binding affinity for a specific DNA than for a non-specific DNA and AdoMet enhances the binding of both substrates but the relative enhancement is greater for specific DNA. The AdoMet analogues S-adenosylhomocysteine (AdoHcy) and sinefungin cannot induce a similar increase in DNA binding although they are bound by the enzyme with approximately the same affinity as AdoMet. The binding affinities for hemimethylated, unmodified and fully modified DNA are all the same indicating that the preference for a hemimethylated substrate has its basis in catalysis rather than in binding. After the methylation reaction AdoHcy becomes the bound product and presumably the decreased DNA binding affinity in the presence of AdoHcy facilitates dissociation (Powell *et al.*, 1993).

AdoMet therefore has a dual role in type I systems. It acts as a methyl donor in catalysis and enhances the affinity of the enzyme for its DNA substrate. Similarly in the *Dam* , M.*Eco* RI and M.*Msp* I enzymes, AdoM*et al*so increases the affinity for DNA (Bergerat and Guschlbauer, 1990; Reich and Mashoon, 1991; Dubey and Roberts, 1992).

In the *Eco* KI methylase, the ternary complex of AdoMet + unmodified DNA + methylase migrates faster in a non-denaturing gel than the binary complex without AdoMet. This indicates that the ternary complex may have a more compact structure. The effect is much less for a hemimethylated DNA substrate and is non-existent for a fully modified DNA. This suggests that although the binding affinities for DNA substrates of different methylation states are very similar, the presence of single methyl groups is sufficient to significantly alter the complex conformation (Powell *et al.*, 1993).

The effect of methylation on DNA structure may in theory be quite considerable. Methylation of the C5 position of cytosine is thought to promote the transition from the B form to the Z form of DNA, particularly when present in longer runs of the CG dinucleotide. The effect of N6 methylation of adenine on DNA structure may also be significant. As a monomer 6-methyladenine exists in two isomeric forms, *cis* and *trans*. The equilibrium is such that the *cis* form is twenty times more favourable than the *trans* form. The *cis* form is unable to base pair with thymine and the disruption to the *cis -trans* equilibrium in order to form such a base pair destabilises the double helix. The presence of the methyl group partly counters this effect by increasing the base-stacking interactions. Nevertheless it may be possible that unpaired regions appear in DNA containing N6 methylated adenine and this could play a significant role in DNA recognition by methylases and restriction endonucleases (Adams and Burdon, 1985). However the crystal structures of the unmethylated and methylated forms of the *Eco* RI recognition site are virtually identical (Fratini *et al.*, 1982; Frederick *et al.*, 1988). In addition, methylation of up to three hundred adenine residues in pBR322 has only an extremely small effect upon DNA structure (Cheng *et al.*, 1985). Further structural studies of DNA containing N6 methylated adenine are therefore needed to determine the extent of any conformational changes which may be theoretically possible.

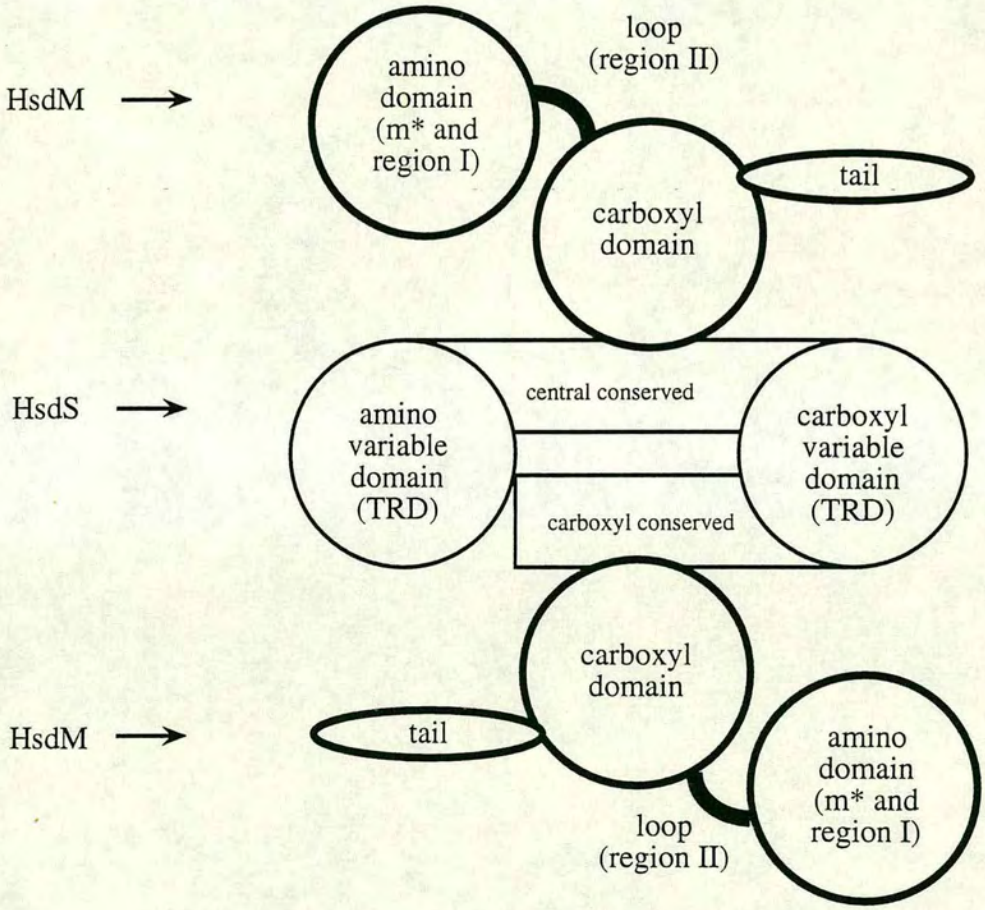
Partial proteolytic digestion of the *Eco* KI methylase has been used to investigate subunit-subunit and subunit-substrate interactions (Cooper and Dryden, 1994). The HsdS subunit comprises two large domains and is protected from proteolysis in the presence of DNA. This is in agreement with the genetic evidence which identifies HsdS as the polypeptide responsible for DNA target specificity. The HsdM subunit also comprises two large domains, joined by a short flexible loop which contains a motif found in all N6-adenine methylases. The presence of

AdoMet provides strong protection to the whole of HsdM apart from a C-terminal tail. The conformation of this tail is sensitive to the methylation state of the DNA in a ternary complex and therefore may help to detect the different methylation states of DNA. The role of the conserved regions of HsdS in subunit-subunit interaction (section 1.3.1) is supported by the finding that the C-terminal domain of HsdM associates with these regions even after the rest of the protein has been digested. A schematic diagram of the *Eco* KI methylase is shown in figure 1.4 (Cooper and Dryden, 1994).

Figure 1.4

A schematic Diagram of the *Eco* KI Methylase. HsdM subunits are shown in heavy lines and HsdS subunits in light lines.

Figure 1.4


HsdM
amino domain (m* and region I)
loop (region II)
carboxyl domain
tail
HsdS
amino variable domain (TRD)
central conserved
carboxyl conserved
carboxyl variable domain (TRD)
HsdM
carboxyl domain
tail
amino domain (m* and region I)
loop (region II)

1.3.4 Mutational Analyses of the *Eco* KI Methylase

The *Eco* KI methylase has a strong preference for a hemimethylated DNA substrate. Mutants have been isolated within *hsdM*, the phenotype of which are consistent with enhanced modification of an unmethylated target (Kelleher *et al.*, 1991). The mutations cluster loosely within a 100 amino acid region of the N terminus. In effect the mutant enzymes are *de novo* methylases (m*), an activity displayed by the wild-type only in the presence of Ral. The mutations may identify part of the HsdM polypeptide involved in recognising target site methylation. This may imply communication between the C-terminal tail and the N terminus of HsdM since it has been demonstrated that the conformational state of the former is sensitive to the DNA methylation state in a ternary complex (section 1.3.3). Furthermore these mutations may provide a molecular explanation for the difference between *de novo* and maintenance methylase activity. The mutants fall into two categories, those which have lost all restriction activity (r^-m^*) and those which retain partial restriction activity ($r^{+/-}$ m*). The mutants were interpreted as forming *Eco* KI complexes and in the latter class restriction and modification may have become competing reactions.

A mutation in *hsdM* has been described which encodes a methylase defective in AdoMet binding (Hubacek, 1973; Bühler and Yuan, 1978). Upon sequencing, the mutation was found to comprise three amino acid changes; P174S, T280S and G388D (Daniel, A., unpublished). Site-directed mutagenesis demonstrated that the single mutation, P174S, is responsible for this phenotype (Barcus, V., unpublished). This proline residue is in a region common to all methyltransferases which use AdoMet as a cofactor (see section 1.4). A second mutation in region I has been described, D173G, the phenotype of which is indicative of no *in vivo* methylation activity (Kelleher, 1990). The sequence of two mutations within *hsdS* have been described. One of these mutations, A204T, is located at the junction of the central conserved region and the carboxyl variable region (Zinkevich *et al.*, 1992). The other mutation, S340F, is located within the carboxyl variable region (Zinkevich *et al.*, 1990). Both mutations have a dominant effect over the wild-type S polypeptide and give the phenotype r^-m^{ts}.

1.4 Conserved Motifs within DNA Methylases

As with any comparison of related proteins the presence of conserved motifs has become more apparent as the number of sequenced genes has increased. The sequence N/D P P F/Y was reported as being characteristic of N6-adenine methylases, being found in M.*Eco* RV, M.*Eco* RI, Dam, and *Eco* KI HsdM, but not in any cytosine methylase or restriction endonuclease (Hattman *et al.*, 1985; Loenen *et al.*, 1987). In type I R-M systems this motif is found in the HsdM polypeptide, consistent with a role in the methylation step. In *Eco* KI HsdM the motif appears as N P P F. A further comparison of sixteen adenine methylases again indicated the prevalence of this motif and it is sometimes termed region II (Klimasauskas *et al.*, 1989, table 1.2). Suggestions for its possible function have included; part of the active site for adenine methylation (Loenen *et al.*, 1987), adenine recognition (Chandrasegaran and Smith, 1988), and AdoMet binding (Lauster *et al.*, 1987; Guschlbauer, 1988).

The importance of this motif to enzyme activity has been demonstrated by mutagenesis of the type IIs methylase *Fok* I (Sugisaki *et al.*, 1989). This enzyme methylates both strands of the asymmetric sequence 5'-GGATG / 3'-CCTAG, to give m6A, and is presumably the result of a fusion between two originally separate methylases. The motif occurs twice in this methylase in the form D P P Y. Mutation of the N-terminal motif (Asp → Gly) gave an enzyme capable of methylating only the 3'- component. Likewise, mutation of the C-terminal motif (Asp → Ala) gave an enzyme capable of methylating only the 5'- component. The double mutant was incapable of methylating either strand.

Support for a role for this motif in the recognition of adenine (or N6-methyladenine) is provided by an observation made on the phage P22 Mnt repressor (Chandrasegaran and Smith, 1988). The first ten amino acids of this repressor, which acts as a homodimer, are A R D D P H F N F R and the DNA operator site is the palindromic sequence ATAGGTCCACGGTGGACCTAT (Vershon *et al.*, 1985). Binding of repressor to operator is reduced by three orders of magnitude if His6 of the operator is mutated to Pro, giving the sequence D P P F. This mutation can be suppressed by a combination of the operator mutation G → A (in both strands of the palindromic sequence) and growth in a dam^+ cell. Since the adenine is part of the sequence GATC, it becomes N6 methylated in a dam^+ environment. This implicates

the interaction between the amino acid sequence D P P F and the modified base N6-methyladenine.

Secondary structure predictions of this motif in four bacterial methylases have indicated that the region forms a pocket of intermediate hydrophobicity, perhaps able to bind AdoMet (Lauster *et al.*, 1987). The two Pro residues are predicted to form a sharp turn in the polypeptide chain and the Phe/Tyr residue could stack with the adenine moiety of AdoMet.

The N4-cytosine methylases contain a very similar motif, S/D P P F/Y (Brooks *et al.*, 1989; Klimasauskas *et al.*, 1989) (table 1.2). The very close resemblance to the motif found in N6-adenine methylases may well reflect the similarity between N4-cytosine methylation and N6-adenine methylation, that is, methylation of an extracyclic amino group as opposed to methylation of cyclic carbon.

Table 1.2

A Sequence Comparison of N4-Cytosine, N6-Adenine and C5-Cytosine DNA Methylases. Region II indicates the sequence similarity between the S/N/D P P F/Y motif of N4-cytosine/N6-adenine methylases and the P C motif of C5-cytosine methylases (adapted from Klimasauskas *et al.*, 1989).

Table 1.2

METHYLASE			Sequence Position	REGION I	Sequence Position	REGION II
Mva I	N4	(α)	69-81	LY DP**F**M**G**S**G** TSLV	267-282	MFD IVV T**S**P**P** **Y**GDSKT
Cfr 9 I	Cytosine	(β)	249-261	VL DP**F**F**G**S**G** TVGV	44-59	SVR CIV T**S**P**P** **Y**WGLRD
Pvu II	···	··	269-281	VV DI**F**G**G**SN TTGL	46-61	SIS LVM T**S**P**P** **F**ALQRK
Eco RV	N6	(α)	35-47	WV EP**F**M**G**T**G** VVAF	186-201	RDD VVY C**D**P**P** **Y**IGRHV
Dpn II	Adenine	··	39-51	YF EP**F**V**G**G**G** ALFF	187-202	TGD FVY V**D**P**P** **Y**IPLSE
T4 dam	···	··	28-40	FV DL**F**C**G**G**L** SVSL	164-179	DGD FVY V**D**P**P** **Y**LITVA
Eco dam	···	··	31-43	LV EP**F**V**G**A**G** SVFL	174-189	DAS VVY C**D**P**P** **Y**APLSA
Hha II	···	(β)	190-202	IL DPAS**G**G**Y** SVFE	30-45	AVK IAF F**D**P**N** **Y**RGVLD
Hinf I	···	··	196-208	VL DP**F**F**G**T**G** TTGA	29-44	SID LIF A**D**P**P** **Y**FMNTD
Eco P1	···	··	432-444	VL DF**F**A**G**S**G** TTAE	116-131	KVM NIY I**D**P**P** **Y**NTGKD
Eco P15	···	··	450-462	IL DF**F**A**G**S**G** TTAH	118-133	KVM NIY I**D**P**P** **Y**NTGKD
Pae R7	···	(γ)	22-34	LL EPSF**G**C**G** DFLL	113-128	QFD FVV G**N**P**P** **Y**VRPEL
Cvi BIII	···	··	53-65	IL EPSC**G**T**G** EIIS	113-128	KFD FIV G**N**P**P** **Y**VVRPS
Eco 571	···	··	38-52	IL EPSC**G**D**G** VFIQ	109-124	IFD GAL G**N**P**P** **F**IRYQF
Taq I	···	··	43-55	VL EPACAH**G** PFLR	98-113	AFD LIL G**N**P**P** **Y**GIVGE
***Eco* KI**	···	··	171-183	VQ DPAAGTA GFLI	259-274	KAH IVA T**N**P**P** **F**GSAAG
*Eco*R124I	···	··	117-129	IY DPAA**G**S**G** SLLL	196-211	PFD AIV S**N**P**P** **Y**SVKWI
Eco EI	···	··	154-166	IM DPAC**G**T**G** GFLA	225-240	QVD VIV T**N**P**P** **F**GGTEE
Pst I	···	··	51-63	IL DAGA**G**V**G** SLTA	145-160	KYN KAI L**N**P**P** **Y**LKIAA
Eco RI	···	?	112-124	SV SS**F**C**G**D**G** DFRS	132-147	KSD IVV T**N**P**P** **F**SLFRE
Hha I	C5		14-26	FI DL**F**A**G**L**G** GFRL	73-90	HD ILC AGF**P**C**Q**A**F**SISGK
Eco RII	Cytosine		98-110	FI DL**F**A**G**I**G** GIRK	177-194	HD VLL AGF**P**C**Q**P**F**SLAGV
Msp I	···		105-117	FI DL**F**S**G**I**G** GIRQ	166-183	HD ILC AGF**P**C**Q**P**F**SHIGK
Dde I	···		3-15	II DL**F**A**G**C**G** GFSH	67-84	VD GII GGP**P**C**Q**A**F**STAGK
Sin I	···		77-89	AL S**F**F**S**G**A**M GLDL	145-162	ID LIM GGP**P**C**Q**A**F**STAGK
Phi 3T	···		6-18	VM SL**F**S**G**I**G** AFEA	69-86	FD LLT SGF**P**C**P**T**F**SVAGA
SPR	···		6-18	VM SL**F**S**G**C**G** GLDL	69-86	FD LLV GGS**P**C**Q**S**F**SVAGH
Bsu RI	···		61-73	VL SL**F**S**G**I**G** AFEA	148-165	CN LIL GGF**P**C**P**G**F**SEAGP
Bsp RI	···		60-72	VL SL**F**C**G**A**G** GLDL	147-164	AN LVI GGF**P**C**P**G**F**SEAGP

Sequence comparison of C5-cytosine methylases shows a series of ten conserved elements (CEs) separated by variable regions (Posfai and Roberts, 1989). Six of these regions are highly conserved; CEs 1, 4, 6, 8, 9 and 10 (Cheng *et al.*, 1993). Separating CEs 8 and 9 is a comparatively long variable region known to be the main determinant in sequence specific DNA recognition (Klimasauskas *et al.*, 1991; Mi and Roberts, 1992). CE 4 has an invariant P C motif which plays a central role in C5-cytosine methylation.

Evidence for the involvement of the P C motif in enzyme activity initially came from studies on thymidylate synthetase, an enzyme which also transfers a methyl group onto the C5 position of pyrimidine (Santi *et al.*, 1984; Dev *et al.*, 1988). The Cys of the P C motif is involved in both substrate binding and in the initiation of catalysis. It was proposed that C5-cytosine methylases function in an analogous manner (Wu and Santi 1987). The catalytic mechanism of these enzymes was shown to involve the attack of the C6 of the substrate cytosine by an enzyme nucleophile and formation of a transient covalent adduct. Adduct formation causes the C5 position of cytosine to become activated and to initiate nucleophilic attack on the methyl group of AdoMet. The methyl group is then transferred to the C5 position and the covalent protein intermediate is released. Sensitivity of the C5 methylase M.*Hha* I to *N* -ethylmaleimide (which attacks sulphydryl groups) and the absolute conservation of the P C motif in C5-cytosine methylases suggested that the Cys residue may act as the enzyme nucleophile. This hypothesis was confirmed using M.*Hae* III as a model for study. Sequencing of DNA-bound peptides following exhaustive proteolysis of a DNA - M.*Hae* III complex led to the identification of the conserved Cys residue as the catalytic nucleophile (Chen *et al.*, 1991).

Site-directed mutagenesis of the P C motif has confirmed its importance to enzyme function (Wilke *et al.,* 1988; Wyszynski *et al.*, 1991; Mi and Roberts, 1993). In the multispecific phage methylase M.SPR, mutation of the Pro residue to either Ser or Leu gave mutants incapable of methylating any of the three target sequences recognised by the wild-type enzyme (Wilke *et al.,* 1988). In M.*Hha* I , mutation of the Cys to Arg, His or Ser gave inactive enzymes. Mutation of Cys to Gly not only led to inactivity but also to cytotoxicity. Biochemical analysis of this mutant indicated abnormally tight DNA binding (Mi and Roberts, 1993). In M.*Eco* RII,

mutation of Cys to a number of amino acids gave inactive enzymes (Wyszynski *et al.*, 1991).

In addition to being invariant in C5-cytosine methylases, the P C motif is also found in all O^6-methylguanine methylases (Rydberg *et al.*, 1990). These enzymes remove a methyl group from the cytotoxic base O^6-methylguanine, restoring guanine to its original state. The Cys residue has been shown to be the methyl acceptor in the human enzyme (von-Wronski, 1991).

In contrast to the C5-cytosine methylases, the N6-adenine and N4-cytosine methylases do not possess a P C motif and this may reflect the different chemistries required to achieve their methylation. The mechanism of the N6-adenine methylase M.*Eco* RI (and presumably other methylases of this class) involves direct methylation of the N6 position of adenine and there is no evidence of a covalent intermediate being formed as in C5-cytosine methylation (Pogolotti *et al.*, 1988). The enzyme-substrate interactions in N6-adenine and N4-cytosine methylation are not yet known.

In an alignment of the polypeptide sequences of the three DNA methylase classes, (table 1.2), a degree of similarity has been discovered between the S/N/D P P F/Y motif of N6-adenine and N4-cytosine methylases and the P C motif of C5-cytosine methylases (Klimasauskas *et al.*, 1989). This would suggest that the role of the two motifs could be analogous, i.e. central in catalysis.

Another motif, common to all DNA methylases, has been noted in several sequence comparisons (Lauster, 1989; Klimasauskas *et al.*, 1989; Ingrosso *et al.*, 1989). This motif is characterised by several positions where hydrophobic residues bracket two or three glycine residues and is often referred to as the F X G X G motif (table 1.2, region I). The range of methylases which possess the motif has been expanded to include not only DNA methylases but also protein, RNA, and bovine hydroxyindole methylases (table 1.3). In addition it is found in an enzyme which has no methylase activity, the human AdoMet decarboxylase. The common factor in all these enzymes is the catalysis of AdoMet. Methylases which use different methyl donors do not contain this motif, for example thymidylate synthetase which uses a tetrahydrofolate derivative (Hardy *et al.*, 1987; Ingrosso *et al.*, 1989).

The N6-adenine and N4-cytosine methylases have been grouped into three classes (α, β, and γ) according to the order and version of the region I and region II motifs (table 1.2). In the α class, region I (F X G X G) comes before region II and the separation of the two regions is over 120 residues. In the γ class, region I (X G X G) comes before region II and their separation is 40-80 residues. In the β class, region II comes before region I. The type I methylases belong to the N6-adenine γ class and in *Eco* KI HsdM the region I motif is Q D P A A G T A G F. The underlined Gly residue is highly conserved in all AdoMet dependent methylases and corresponds to the middle glycine of the F X G X G motif.

Table 1.3

A Sequence Comparison indicating Similarity between DNA Methylases, RNA Methylases, Small Molecule Methylases and AdoMet Decarboxylase (adapted from Ingrosso *et al.*, 1989).

Table 1.3

ENZYME (MT=methylase)	Sequence Position	REGION I
D-Asp / L-isoAsp MT	81-97	AL DVGS**G**S**G** ILTACFAR
Hydroxyindole MT	182-198	IC DLGG**G**S**G** ALAKACVS
Rat glycine MT	255-271	VQ EAFG**G**RC QHSVAFTV
Rat catechol *O* -MT	61-76	VL ELGAYC**G** SAVRMARL
erm F RNA MT	37-53	VL DIGA**G**K**G** FLTVHLLK
erm S RNA MT	33-49	VY EIGT**G**K**G** HLTTKLAK
erm C RNA MT	34-50	IF EIGS**G**K**G** HFTLELVK
erm A' RNA MT	61-77	VV EAGP**G**E**G** LLTRELAR
erm D RNA MT	48-64	VL ELGA**G**K**G** ALTTVLSQ
erm G RNA MT	34-50	IF EIGA**G**K**G** HFTAELVK
erm Y RNA MT	34-50	IF EIGA**G**K**G** HFTAELVK
erm X RNA MT	33-49	VY EIGT**G**K**G** HLTTKLAK
erm E RNA MT	65-81	VL EAGP**G**E**G** LLTRELAD
Pvu II DNA MT (N4)	269-285	VV DIFG**G**SN TTGLVAER
Pae R7 DNA MT (N6)	22-38	LL EPSF**G**G**G** DFLLPIIQ
Dpn II DNA MT ··	39-55	YF EPFV**G**G**G** ALFFDLAP
Dpn A DNA MT ··	211-227	IL DPFV**G**S**G** TTGVVAKR
Dam DNA MT ··	31-47	LV EPFV**G**A**G** SVFLNTDF
Pst I DNA MT ··	61-77	IL DAGA**G**V**G** SLTAAFVQ
***Eco* KI DNA MT** ··	171-187	VQ DPAA**G**TA GFLIEADR
Hha I DNA MT (C5)	14-30	FI DLFA**G**L**G** GFRLALES
SPR DNA MT ··	6-22	VM SLFS**G**I**G** AFEAALRN
Phi3T DNA MT ··	6-22	VM SLFS**G**I**G** AFEAALRN
Bsu RI DNA MT ··	61-77	VL SLFS**G**C**G** GLDLGFEL
Bsp RI DNA MT ··	60-76	VL SLFC**G**A**G** GLDLGFEL
Mouse DNA MT	1022-1038	TL DVFS**G**C**G** GLSEGFHQ
AdoMet decarboxylase	22-38	QP DANQ**G**S**G** DLRTIPRS

1.5 DNA Methylase Crystal Structure

The only crystal structure of a DNA methylase is that of the type II C5-cytosine methylase M.*Hha* I (figure 1.5). Both the enzyme-AdoMet and enzyme-AdoHcy-DNA structures have now been solved (Cheng *et al.*, 1993; Klimasauskas *et al.*, 1994).

The structures demonstrate the interactions of, among others, the P C and F X G X G motifs discussed in section 1.4 and the core of the structure is dominated by the ten conserved elements of C5-cytosine methylases. The enzyme comprises a large amino domain connected to a small carboxy domain via a hinge region. These three sections correspond to the three sides of a cleft large enough to accommodate a DNA substrate. A cavity in the large domain face of the cleft contains the AdoMet-binding site, the adenosyl moiety extending into the cleft. The binding pocket involves residues from the N terminus and residues from the C terminus. As predicted from sequence comparisons, the region I motif, F X G X G, contributes significantly to AdoMet binding. The amino acids surrounding this motif form a 'β sheet - α helix - β sheet' structure, the two Gly residues forming a tight loop between the first β sheet and the α helix. These residues appear to be crucial in positioning the adenine ring of AdoMet in its correct conformation to allow close contacts with the protein main chain. The Phe residue interacts hydrophobically with AdoMet by contributing to a platform on one side of the purine and ribose rings. The Pro residue of the P C motif also contributes to this platform by interacting with the other side of these rings. The AdoMet-binding site bears no similarity to the AdoMet-binding pocket of the MetJ repressor which uses AdoM*et al*losterically rather than catalytically (Rafferty *et al.*, 1989; Somers and Philips, 1992). However the structure does resemble the Rossmann fold found within the NAD-binding motif which also comprises a β–α –β structure containing a glycine-rich turn.

In the enzyme-AdoMet binary complex the carbon atom of the AdoMet methyl group is nearly 10Å away from the sulphur atom of the catalytic Cys residue of the PC motif. The reaction mechanism is known to involve the formation of a transient covalent link between the cysteine residue and the target cytosine (section 1.4). After allowing for the pyrimidine ring of cytosine the distance between the sulphur atom of Cys and the cytosine C6 position is larger than would be expected for covalent bond formation. In addition the distance between the cytosine C5

position and the donor methyl group is also greater than would be expected for catalysis.

Determination of the tertiary structure accounts for these large separations between the atoms involved within the active site (Klimasauskas *et al.*, 1994). None of the usual DNA recognising structural motifs are present in the structure. The complex shows a quite unanticipated arrangement in which the target cytosine is rotated out of the DNA helix and into the catalytic site. The cytosine is held in place by being buried within the cleft and is stabilised by making contacts with four residues in CEs 4, 5 and 8. Three of these residues are conserved across all C5-cytosine methylases. The vacancy left by the exiting cytosine is taken up by a Gln residue which H-bonds to the remaining lone guanine. During the DNA binding process there is a conformational change in the protein which involves the amino domain loop region containing the P C motif moving nearly 25Å, taking it into the DNA minor groove and positioning the Cys residue for catalysis. Two more loops, from the carboxy domain, interact with the DNA helix from the major groove side, contacting non-overlapping portions of the recognition sequence in a diagonal arrangement.

Given the pattern of conserved elements found in all C5-cytosine methylases it is highly probable that the interactions discussed above are readily applicable to other methylases of this class. However the degree of similarity of these interactions to those of N6-adenine (and N4-cytosine) methylases remains to be elucidated.

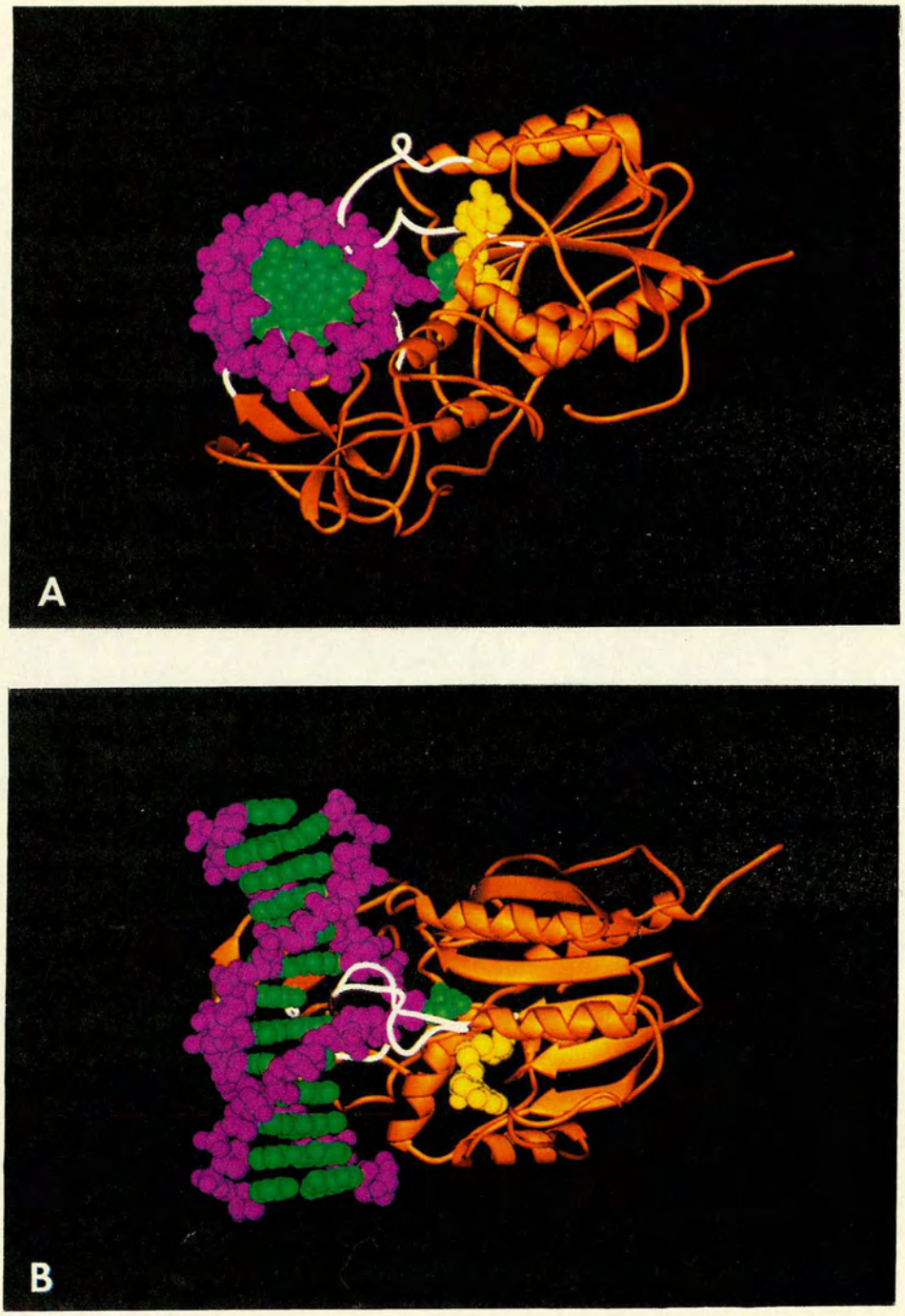
The work in this thesis describes a site-directed mutagenesis approach to investigate the role of the N/D P P F/Y and F X G X G motifs as found in the HsdM polypeptide of *Eco* KI. *In vivo* and *in vitro* techniques have been used to determine the effect of mutation in these regions of the methylase upon activity, stability, AdoMet binding and DNA binding.

Figure 1.5

Graphic Representation of the Complex of M.Hha I covalently bound to a 13-mer DNA Duplex containing its Recognition Sequence. The end product of the reaction, AdoHcy, is also present (yellow). The protein is in brown, the DNA backbone is in magenta and the DNA bases are in green. The active site loop (cotaining the P C motif) and the two recognition loops are in white (Klimasauskas *et al.*, 1994).

A) View looking down the DNA helix axis, in which the DNA can be seen to be between the large domain (on the right) and the small domain (on the left).
B) Side view from the minor groove.

Figure 1.5



Chapter Two

MATERIALS AND METHODS

2.1 STRAINS

Table 2.1 Bacterial Strains

Strain number	Relevant Features	Reference/Source
NM522	(*lac -pro*)Δ, *hsd MS* Δ5, F' (*lac Z* ΔM15, *lac I* ^{q})	Gough and Murray, 1983
NM679	*hsd RMS* Δ	N. E. Murray
NM717	*mutL*, *hsd RMS* Δ	N. E. Murray (Zoller and Smith, 1983)
C600	*hsd* K^{+}	Appleyard, 1954

Table 2.2 Phage Strains

Strain Number	Relevant Features	Reference/Source
NM 1065	λ*Bam* 1°, *hsd MΔ S* $^{+}$	Sain and Murray, 1980
NM 1048	λ *hsd MS* $^{+}$, cI857	Sain and Murray, 1980
M13 mp19	vector for DNA sequencing	Yanish-Perron *et al.*, 1985

Table 2.3 Plasmids

Plasmid	Relevant Features	Reference/Source
pJF118HE	overexpression vector using tac promoter	Furst *et al.*, 1986

2.2 ENZYMES AND CHEMICALS

T4 DNA ligase, dNTPs and ddNTPs were from Pharmacia; Sequenase enzyme from USB; DNA polymerase and restriction enzymes from New England Biolabs (NEB); DNase I, RNase A and lysozyme from Sigma Chemical Company Ltd. T4 polynucleotide kinase (T4 PNK) from S. Bruce (Institute of Cell and Molecular Biology, University of Edinburgh).

Ampicillin was from Beecham Pharmaceuticals; Vitamin B_1, DTT, β-mercaptoethanol and IPTG from Sigma; X-gal from Boehringer.
M13 sequencing primer (17-mer) was from NEB; all other synthetic oligonucleotide primers from Oswel DNA Service (Edinburgh).
Nitrocellulose filters were from Schleicher and Schuell; autoradiographic film from Amersham International.
Radiolabelled compounds; [α-^{32}P]dCTP, [α-^{35}S]dATP, [γ-^{32}P]ATP and [*methyl*-^{3}H]AdoMet were from Amersham International. Scintillation fluid from National Diagnostics. Acrylamide and *N,N* -methylene bis-acrylamide were from BDH and NBL.
Adomet was a gift from New England Biolabs or from Sigma. HPLC purified ANS from Molecular probes. Coomassie blue R250, MES, tris (hydroxymethyl) amino-methane (Tris) and SDS were from Sigma. Chromatography media from Pharmacia. Guanidine chloride (Ultrapure) from Bethesda Research Laboratories.
Standard laboratory chemicals were from Sigma, BDH or Fisons.

2.3 MEDIA

L-Broth: Difco Bacto tryptone (10g), Difco Bacto yeast extract (5g), NaCl (5g), distilled H_2O to 1 litre; adjusted to pH 7.2 with NaOH before autoclaving.
L-agar: Difco Bacto tryptone (10g), Difco Bacto yeast extract (5g), NaCl (5g), Difco agar (15g), distilled H_2O to 1 litre; adjusted to pH 7.2 with NaOH before autoclaving.
BBL-Agar: Baltimore Biological Labs. Trypticase (10g), NaCl (5g), Difco agar (10g), distilled H_2O to 1 litre.
BBL-Top Agar: As for BBL-agar but only 6.5g Difco agar added per litre.
Minimal Agar: Difco agar (4g), distilled H_2O to 300ml. After autoclaving the following sterile solutions were added: 5x spizizen salts (80ml), glucose (4ml, 20% w/v), vitamin B_1 (0.1ml, 2mg/ml).
2x TY Broth: Difco Bacto tryptone (16g), Difco Bacto yeast extract (10g), NaCl (10g), distilled H_2O to 1 litre.
5x Spizizen Salts: $(NH_4)_2SO_4$ (10g), K_2HPO_4 (70g), KH_2PO_4 (30g), tri-sodium citrate dihydrate (30g), $MgSO_4.7H_2O$ (1g), distilled H_2O to 1 litre.
Phage Buffer: Na_2HPO_4 (7g), KH_2PO_4 (3g), NaCl (5g) , $MgSO_4.7H_2O$ (10ml, 0.01M), $CaCl_2$ (10ml, 0.01M), gelatin (1ml, 1% w/v), distilled H_2O to 1 litre.

Antibiotics: Ampicillin was used at a concentration of 100μg/ml. When used in plates, the antibiotic was added to molten agar immediately prior to pouring.
X-gal Indicator plates: X-gal (20μl, 20mg/ml), IPTG (20μl, 20mg/ml), per 2.5ml of BBL top agar.

All media were sterilised by autoclaving (15lb/in^2, 15 minutes).

2.4 STANDARD SOLUTIONS

TE Buffer: Tris (10mM), EDTA (1mM) adjusted to appropriate pH with HCl.
20x SSC: NaCl (3M), tri-sodium citrate (0.3M).
20x TBE Buffer: Tris (1.78M), boric acid (1.78M), EDTA (50mM).
Ethidium Bromide: 10mg/ml in distilled H_2O. Stored in dark at 4°C.

2.5 MICROBIAL TECHNIQUES

2.5.1 Preparation of Plating Cells

1) A fresh overnight culture was diluted in L-Broth (20x) and grown at the required temperature to mid log-phase.

2) The cells were harvested by centrifugation (3000g, 5 minutes).

3) The cells were resuspended in half the original volume of Mg_2SO_4 (10mM) and stored at 4°C.

2.5.2 Preparation of Competent Cells

1) A fresh overnight culture was diluted in L-broth (50x) and grown at 37°C to an $O.D._{650}$ of 0.7.

2) The cells were harvested by centrifugation (3000g, 5 minutes).

3) The cells were resuspended in an equal volume of ice-cold $MgCl_2$ (100mM).

4) The cells were harvested again and resuspended in one half the original volume of ice-cold $MgCl_2$ (100mM).

5) The cells were harvested for a third time and resuspended in one tenth the original volume of ice-cold $CaCl_2$ (100mM).

2.5.3 Preparation of λ Phage; Plate lysates

1) A single plaque was picked into phage buffer (1ml, containing a drop of chloroform) and vortexed.

2) Phage suspension (0.1ml) was added to fresh plating cells (0.2ml) and the phage left to adsorb for 15 minutes.

3) BBL-top agar (3ml) was added to the mixture and poured onto a fresh L-agar plate.

4) The plate was incubated (not inverted) at the required temperature until confluent lysis was observed, typically 6-8 hours. A control of cells without phage provided a good comparison.

5) L-broth (5ml) was added and the plate stored at 4°C overnight.

6) The L-broth was decanted, one drop of chloroform added and the lysate was vortexed.

7) The supernate, containing phage, was harvested (3000g, 10 minutes) and titred before storing at 4°C.

2.5.4 Phage Titration

1) Serial dilutions of phage stock were prepared using phage buffer.

2) 0.1ml of each dilution was added to plating cells (0.2ml) and left to adsorb for 15 minutes.

3) The mixture was plated out in BBL-top agar (3ml) onto BBL plates and incubated overnight at the required temperature.

2.5.5 Spot Tests

1) Plating cells (0.2ml) were spread onto a BBL plate in BBL top agar (3ml).

2) 10µl spots of phage dilutions (made with phage buffer) were applied to the bacterial lawn and allowed to dry before overnight incubation at the required temperature.

2.6 DNA TECHNIQUES

2.6.1 Precipitation of DNA with Ethanol and Sodium Acetate

1) DNA in solution was precipitated by adding 0.1x volume of sodium acetate (3M, pH 5.0) and 2.5x volume of ethanol. The mixture was then incubated at -70°C for 30 minutes.

2) The precipitate was harvested (11600g, 10 minutes) and washed with ethanol (70% w/v).

3) The pellet was dried under vacuum and resuspend in the required volume of TE buffer (pH 8.0).

2.6.2 Alkaline Lysis Plasmid Miniprep

An adaptation of Birnboim and Doly (1979).

1) A fresh overnight culture (1.5ml) carrying the plasmid was harvested by centrifugation (11600g, 5 minutes).

2) The pellet was resuspended in 100µl lysis solution (Tris-HCl, 25mM; EDTA, 10mM; glucose, 0.9%) and incubated at room temperature for 10 minutes.

3) 200µl alkaline SDS (NaOH, 0.2M; SDS, 1%) was added and mixed by gentle inversion, followed by incubation on ice for 5 minutes.

4) Chromosomal DNA was precipitated by adding sodium acetate (150 µl, 3M, pH 5.0) and incubation on ice for 30 minutes. The supernate was harvested by centrifugation (11600g, 15 minutes).

5) In order to precipitate plasmid DNA, ethanol (2.5x volume, -20°C) was added to the supernate and after mixing, incubated at -70°C for 30 minutes .

6) The precipitated plasmid DNA was harvested by centrifugation (11600g, 15 minutes) and washed with ethanol (70%,w/v).

7) The DNA pellet was dried under vacuum and resuspended in TE buffer (20µl, pH 8.0).

2.6.3 Preparation of M13 Replicative Form (RF) DNA

1) Early log-phase cells were prepared by diluting a fresh overnight culture in L-broth (100x) and growing at 37°C to an $O.D._{650}$ of 0.2.

2) A single M13 plaque was picked into cells (1.5ml) and growth continued at 37°C for a further 6-8 hours.

3) The cells were harvested by centrifugation (11600g, 5 minutes) and the resulting lysate titred (approximately 10^{11} pfu/ml) and used to infect a culture of early log-phase cells (50ml) at a final concentration of 10^9 pfu/ml.

4) The culture was grown overnight at 37°C and the phage supernate was harvested (3000g, 5 minutes) and its titre determined.

5) A culture (500ml) of late log-phase cells was infected with phage at 10^{11} pfu/ml and growth continued for 2 hours at 37°C.

6) The cells were harvested by centrifugation (4000g, 5 minutes) and after resuspension in lysis solution (3.5ml, see section 2.6.2) were incubated on ice for 20 minutes.

7) Lysozyme (8mg, dissolved in 0.5ml lysis solution) was added to the resuspended cells and incubated on ice for 20 minutes.

8) Alkaline SDS solution (8ml, see section 2.6.2) was added and incubation continued for a further 10 minutes.

9) In order to precipitate unwanted protein and chromosomal DNA, sodium acetate (5ml, 3M, pH 5.0) was added followed by gentle stirring and incubation on ice for 10 minutes.

10) The supernate containing the M13 RF DNA was harvested by centrifugation at 4°C (12000g, 15 minutes).

11) In order to remove any remaining protein, the supernate was extracted with phenol/chloroform (1:1).

12) DNA was precipitated with ethanol and sodium acetate and resuspended in TE buffer (500µl, pH 8.0).

13) RNase A (5µl, 10mg/ml) was added to the resuspended DNA and incubated at

37°C for 20 minutes.

14) Protein was removed by repeated phenol/chloroform (1:1) extractions until the interface was clear.

15) The resulting plasmid DNA was precipitated with ethanol and sodium acetate and resuspended in TE buffer (500µl, pH 8.0).

2.6.4 Preparation of CsCl purified λ Phage DNA

1) A fresh overnight bacterial culture was diluted 50x in L-broth (100ml), supplemented with $MgSO_4$ (10mM). Growth was continued at 37°C until an $O.D._{650}$ was reached.

2) Phage was added to the cells (to a m.o.i. of 0.2) and incubation continued at 37°C whilst periodically monitoring turbidity in order to establish when lysis had occurred (usually 3-4 hours).

3) Chloroform (0.2ml) was added and shaking continued for a further 15 minutes.

4) The phage supernate was harvested (4000g, 5 minutes).

5) To the phage was added; 4% w/v NaCl, 1µg/ml DNase I and 1µg/ml RNase A, and allowed to stand at room temperature for 1 hour.

6) Phage were precipitated by adding 10g polyethylene glycol (PEG 6000) and precipitation was allowed to go to completion by gentle swirling at 4°C overnight.

7) The phage precipitate was recovered by centrifugation (16000g, 10 minutes) and resuspended in phage buffer (10ml) by gently swirling at 4°C for 2-3 hours.

8) The phage supernate was harvested (4000g, 5 minutes).

9) Phage were concentrated by density centrifugation using CsCl step gradients. The CsCl solutions were made using phage buffer and clarified by centrifugation if necessary. CsCl solution (1.5ml, 31% w/w) was pipetted into a 14ml polycarbonate tube and using a pasteur pipette, underlaid with a second CsCl solution (1ml, 45% w/w). Finally these two steps were underlaid with a third CsCl solution (1ml, 56% w/w).

10) Phage solution was carefully added to the gradient and centrifuged at 20°C (140000g, 2 hours).

11) The phage band was collected by piercing the tube using a 21 gauge needle and a syringe.

12) CsCl was removed from the phage by overnight dialysis against TE buffer (pH 8.0) at 4°C.

13) In order to remove protein, the phage were extracted three times with phenol (pre-equilibrated with TE).

14) Phenol was removed by dialysis against TE at 4°C for 24 hours with several buffer changes.

2.6.5 Restriction Endonuclease Digestion of DNA

Restriction endonuclease digestion of DNA was carried out in volumes of 10-30µl containing 0.2-0.8µg DNA under conditions specified by the suppliers.

2.6.6 Ligation of DNA

Ligation of DNA using T4 DNA ligase was carried out in volumes of 10-20µl containing 0.2-0.4 µg DNA under conditions specified by the suppliers (incubated at 16°C overnight and terminated by heat inactivation at 70°C for 10 minutes).

2.6.7 Agarose Gel Electrophoresis

For both analysis of DNA (standard agarose) and purification of DNA (low melting-point agarose), horizontal submerged gel electrophoresis equipment was used. Agarose gels were made in the concentration range 0.65-1.0% w/v, depending on anticipated DNA fragment size. A 1xTBE buffer system was used. DNA samples (0.2-0.5µg) were applied in combination with Ficoll loading dye (20% w/v in H_2O, with bromophenol blue) and electrophoresis carried out at 10V/cm. Following staining with ethidium bromide, DNA was visualised under long-wave UV light.

2.6.8 DNA Fragment Preparation from Low Melting-Point Agarose Gels

1) The DNA band was cut from the gel and added to 1x volume TBE, NaCl (0.2M) followed by melting in a 65°C waterbath.

2) Protein was removed by extracting twice with phenol (pre-equilibrated with 1xTBE, 0.1M NaCl).

3) In order to get rid of ethidium bromide the aqueous layer was extracted thrice with 4x volume butanol.

4) DNA was precipitated with ethanol and sodium acetate, and the resulting pellet washed twice with ethanol (70%, v/v) followed by drying under vacuum and resuspension in H_2O.

2.6.9 Transfection and Transformation of Competent Cells

1) DNA was added to competent cells and incubated on ice for 30 minutes.

2) The cells were incubated at 42°C for 1.5 minutes (heat-shock) and returned to ice for a further 30 minutes.

3) For transfection, the cells were plated in BBL-top agar (2.5ml) onto BBL plates and incubated at 37°C overnight. For transformation, the cells were added to L-broth (1ml) and the culture incubated at 37°C for 1 hour. Aliquots of transformed cells were spread on L-agar plates and incubated at 37°C overnight.

2.6.10 *In vitro* Packaging of λ DNA

In vitro packaging was performed using mixes kindly donated by Annette Campbell. The reaction mixture was prepared by adding the following components in the order listed.

Component	Volume
1)Buffer A*	7µl
2)DNA	0.4-1µg (in a volume of 7-10µl)
3)Buffer M1	1µl
4)Sonicated extract	6µl
5)Freeze thaw lysate	10µl

Buffer A: Tris-HCl (20µl, 1M, pH 8.0); EDTA (2µl, 0.5M, pH 7.5); $MgCl_2$ (3µl, 1M); H_2O (975µl).

Buffer M1: Tris-HCl (6µl, 0.5M, pH 7.5); putrescene/spermidine (300µl, 0.1M/50mM); $MgCl_2$ (9µl, 1M); ATP (75µl, 0.1M); β-mercaptoethanol (1µl); H_2O (110µl).

The reaction was incubated at 23°C for 2 hours, followed by the addition of phage buffer (500 µl) and titration to assess packaging efficiency.

2.6.11 Transfer of DNA from Plaques to Nitrocellulose

An adaptation of Benton and Davies (1977).

1) Phage were plated in BBL-top agar on dry plates (BBL-agar for λ, minimal agar for M13) at a density of approximately 300 pfu/ plate.

2) After overnight incubation at the appropriate temperature the plates were cooled by storage at 4°C for 1 hour.

3) Nitrocellulose filter discs were laid onto the agar surface and left for 1 minute to allow DNA transfer. After marking the discs to allow later orientation with respect to the plate, they were removed from the plates and laid plaque side uppermost for 2 minutes on a piece of blotting paper saturated with denaturation buffer (NaOH,

0.5M; NaCl, 1.5M).

4) The filters were placed in neutralisation buffer (Tris-HCl, 0.5M, pH 7.4; NaCl, 3M) for 3 minutes, rinsed in 2x SSC and blotted dry before baking at 80°C under vacuum for 2 hours.

2.6.12 Radiolabelling of Double-Stranded Probes by Nick-Translation and Hybridisation to Filters.

An adaptation of the methods described by Maniatis *et al.*, (1975) and Rigby *et al.*, (1977).

1) [α-^{32}P]dCTP (10µCi) was added to dNTP buffer (20µl), DNase I (1µl, 2x 10^5 mg/ml), DNA (0.5-1.0µg in 2 µl), DNA polymerase I (0.5 units), and the reaction incubated at 15°C for 2 hours followed by termination with EDTA (100µl, 10mM, pH 8.0).

2) In order to remove unincorporated radiolabel the completed reaction was passed down a column of Sephadex G-50 equilibrated with TE buffer. The DNA was collected as the first peak of radioactivity in a volume of 0.5-1.0ml.

3) Nitrocellulose filters (bearing test and control DNA) were pre-hybridised in hybridisation buffer (50ml) for 1 hour at 37°C. Meanwhile the radiolabelled DNA probe was added to non-specific DNA (calf thymus, 250µg), denatured by boiling for 10 minutes and immediately placed on ice to prevent renaturation.

4) The probe was added to the filters in a sealed polythene bag in hybridisation buffer (10-20ml) and hybridisation continued overnight at 37°C with gentle shaking.

5) Filters were washed twice in 2x SSC, SDS (0.1%, w/v), for 30 minutes at room temperature, followed by rinsing in 1x SSC and blotted dry.

6) Hybridisation of probe to filters was detected by autoradiography at -70°C using intensifying screens.

1x dNTP buffer:	nick translation buffer (250µl); dATP/dTTP/dGTP (10µl, 2mM); β-mercaptoethanol (25µl); distilled H_2O (737µl).
Nick Translation Buffer:	Tris-HCl (210mM, pH 7.5); $MgCl_2$ (21mM); BSA (20µg/ml).
Hybridisation Buffer:	formamide (50% v/v); 4x SSC; 1x Denhardt's solution.
20x Denhardt's Solution:	polyvinylpyrolidine (0.4% w/v); Ficoll 400 (0.4% w/v); BSA (0.4% w/v).

2.6.13 End Labelling of Oligonucleotides

End-labelled oligonucleotides were used in site-directed mutagenesis and gel retardation experiments.

1) The following components were added together:

Component	Volume
[γ-^{32}P]ATP(20µCi)	2µl
Tris (1M pH 7.5)	1.5µl
$MgCl_2$ (100mM)	3µl,
DTT (100mM)	1.5µl,
oligonucleotide	1µl, (~0.1nmoles DNA)

2) T4 PNK (4-5 units) was added and the mixture incubated at 37°C for 45 minutes.

3) The reaction was stopped by heating at 65°C for 10 minutes.

2.7 DNA SEQUENCING (Sanger *et al.*, 1977 and 1980)

2.7.1. Preparation of Single-Stranded DNA Templates from M13 Lysates

1) A purified M13 plaque was used to infect a fresh exponential culture (1.5ml) of host cells diluted (100x) in 2x TY medium. The culture was incubated with aeration at 37°C for 5 hours and the unwanted cells discarded (11600g, 5 minutes).

2) The supernatant phage were precipitated overnight by addition of 200µl PEG/NaCl solution (PEG 6000, 20%w/v; NaCl, 2.5M) and incubation at 4°C.

3) Phage were harvested by centrifugation (11600g, 5 minutes) and all traces of PEG removed with a drawn-out pasteur pipette. The pellet was resuspended in TE buffer (100µl, pH 8.0) and extracted with an equal volume of phenol (TE-equilibrated).

4) Phage DNA was precipitated with ethanol and sodium acetate and then resuspended in TE buffer (50µl, pH 8.0) and stored at -20°C.

2.7.2 Annealing of Oligonucleotide Primer to Template DNA

1) The following components were combined:

Component	Volume
Primer	1µl
TM buffer	1µl
DNA	8µl (approx. 1-2µg)

TM buffer: Tris (100mM), $MgCl_2$ (50mM); adjusted to pH 8.5 with HCl.

2) The mixture was incubated at 60°C for 1 hour, then allowed to cool to room temperature.

2.7.3 Dideoxy Chain Termination Sequencing Reactions with Sequenase Enzyme

A) Labelling reaction

1) The following components were added in the order listed:

Component	Volume
Template/primer	10µl
DTT (100mM)	1µl
Labelling mix	2µl
[α-^{35}S]dATP	0.5µl
Sequenase	2µl

Labelling mix: dCTP (1.5mM), dGTP (1.5mM), dTTP (1.5mM).
After thorough mixing, the tube was incubated at room temperature for 5 minutes.

B) Termination reaction

1) Four Eppendorf tubes (labelled G, A, T, C) each containing 2.5µl of the appropriate termination mix (see table 2.7) were pre-warmed to 37°C.

2) When the labelling reaction was complete, 3.5µl was transferred to each of the four tubes containing the termination mixes.

3) After very brief centrifugation in order to mix the components the contents were

incubated at 37°C in a water bath for 5 minutes.

4) The reactions were stopped by the addition of formamide dye (4µl) and stored at -20°C until required.

5) Before loading, the samples were incubated in a bath of boiling water for 5 minutes.

Table 2.7 Termination Mixes

Component	T MIX (µl)	C MIX (µl)	G MIX (µl)	A MIX (µl)
0.5mM dTTP	32	32	32	32
0.5mM dCTP	32	32	32	32
0.5mM dGTP	32	32	32	32
0.5mM dATP	32	32	32	32
Distilled H_2O	52	52	52	52
0.5M NaCl	20	20	20	20
1mM ddTTP	1.6	-	-	-
1mM ddCTP	-	1.6	-	-
1mM ddGTP	-	-	1.6	-
1mM ddATP	-	-	-	1.6

2.7.4 DNA Sequencing Gels

1) Gel plates were cleaned with ethanol before being taped together, separated by a 0.4mm Plastikard spacer. The bound plates were clamped tight with bulldog clips.

2) A crude buffer gradient gel (Biggin *et al.*, 1983) allowed at least 250 bases to be read from a single priming reaction. TBE gel mix (2.5x, 7ml) was added to ammonium persulphate (AMPS) (14µl, 25%w/v) and TEMED (35µl). In a separate vessel the following were mixed: TBE gel mix (0.5x, 40ml), (AMPS) (70µl, 25%w/v) and TEMED (35µl).

3) Using a 10ml pipette, TBE gel mix (2.5x, 4ml) and TBE gel mix (0.5x, 6ml) were taken up and allowed to form a crude gradient by allowing a few bubbles to pass up the pipette. The mixture was poured between the clamped plates.

4) The remaining TBE mix (2.5x) was used to top up the space left in order to form a full length gel.

5) Loading wells were formed using sharkstooth combs.

6) Samples were loaded onto gels using a flat-ended Gilson tip and gels run at 25W in 1xTBE buffer for approximately 2.5 hours (until the bromophenol blue dye was

within 3cm of the bottom of the gel).

7) After electrophoresis, gels were fixed in 10% methanol/10% acetic acid for 20 minutes.

8) Gels were transferred to blotting paper, covered with Saranwrap plastic film and dried on a vacuum drier at 80°C. The Saranwrap was removed before overnight exposure to autradiographic film.

40% acrylamide solution: acrylamide (38g), N,N'-methylene bisacrylamide (2g); made up to 100ml with distilled H_2O and deionised by stirring with Amberlite MB-1 resin (5g); filtered before storage at 4°C in dark.

0.5x TBE gel mix: urea (92g), 40% acrylamide solution (30ml), 10x TBE (10ml); made up to 200ml in distilled H_2O and stored at 4°C in dark.

2.5x TBE gel mix: urea (46g), 40% acrylamide solution (15ml), 10x TBE (25ml), sucrose (5g) bromophenol blue (5mg); made up to 100ml in distilled H_2O and stored at 4°C in dark.

2.8 SITE-DIRECTED MUTAGENESIS (Zoller and Smith, 1983)

This method uses two primers to increase the yield of double stranded molecules following the extension reaction. A *mutL* strain (NM717) is used for the recovery of the phage since it provides an enhanced recovery of mutated molecules as a result of failure to carry out normal mismatch repair.

2.8.1 Annealing of Primers to Template DNA

Single-stranded M13 template DNA containing the fragment of interest was prepared as in section 2.7.1. The mutagenic and universal primers were end labelled as in section 2.6.13.

1) The following components were added together:

Component	Volume
Template DNA	7µl, 10µg
End labelled mutagenic oligonucleotide	1µl, ~ 8pmoles
End labelled universal oligonucleotide	1µl, ~ 8pmoles
Solution A	1µl

Solution A: Tris (0.2M, pH 7.5), $MgCl_2$ (0.1M), NaCl (0.5M), DTT (10mM).

2) The mixture was heated at 90-100°C for 3 minutes and cooled slowly to room temperature.

2.8.2 Extension

1) Solution C (10µl) was added to the annealed DNA.

2) Klenow enzyme (2-3 units) was added and the extension reaction incubated overnight at 15°C.

Solution C: Solution B(1µl), dCTP (1µl, 10mM), dGTP (1µl, 10mM), dTTP (1µl, 10mM), dATP (1µl, 10mM), ATP (1µl, 10mM), DTT (1µl, 100mM), H_2O (2µl).

Solution B: Tris (0.2M, pH 7.5), $MgCl_2$ (0.1M).

2.8.3 Transfection

1) *mutL* competent cells were prepared as in section 2.5.2 and transfected using the theextended DNA from 2.8.2.

2.8.4 Isolation of M13 Phage carrying the required Mutation

1) M13 plaques from the *mutL* transfection were purified in NM522 and prepared for screening by plating at a density of ~ 300 pfu/plate.

2) DNA was transferred from plaques to nitrocellulose filters as in section 2.6.11. The filters were prehybridised (6x SSC, 5x Denhardt's solution, 0.1% SDS) for 1 hour and then hybridised overnight using ^{32}P end labelled mutagenic oligonucleotide as probe.

3) Differential screening was used to isolate those phage carrying the required mutation. Filters were initially washed (6x SSC) at 37°C and then at progressively higher temperatures ; 45°C, 52°C, 56°C, 60°C. After each wash an autoradiographic image was obtained. Potential positives were isolated on the basis of their stronger hybridisation.

4) After purification on NM522, potential positives were sequenced (section 2.7) to determine if they carried the required mutation. Genuine positives were sequenced fully (i.e. entire fragment contained by the M13 vector) in order to identify any unwanted mutations.

2.9 PROTEIN TECHNIQUES

2.9.1 Protein Purification and Storage (Dryden *et al.*, 1993)

1) Freshly transformed cells (NM679) were grown in 2.4 litres of L broth (using 6x 2 litre baffled flasks to increase aeration) containing ampicillin (30μM). All mutants were grown at 37°C except G177D which was grown at 25°C. Expression was induced at an $O.D._{650}$ of 0.5 by adding IPTG to 1mM. Growth was continued for 4 hours and cells were then harvested by centrifugation (4000g, 15 minutes) and the cell paste (typically 6-7g) was stored at -70°C until required. All subsequent steps in the purification were performed at 4°C.

2) The buffer used throughout the purification procedure (buffer A) was Tris (20mM), MES (20mM), $MgCl_2$ (10mM), β-mercaptoethanol (7mM), EDTA (0.1mM), adjusted to pH 6.5 or 8.0 with HCl or NaOH. The cell free extract of soluble proteins was prepared by resuspending the cell paste in buffer A (20ml, pH 6.5) and sonicating, on ice, in 30 second bursts for a total time in minutes equal to the weight of cell paste. In order to reduce the effect of any protease action, the sonication was done in the presence of two inhibitors, benzamidine (20mM) and phenylmethylsulphonyl fluoride (10mM).

3) The cell extract was clarified by centrifugation (30000g, 3 hours) and the supernatant applied to a CM-Sepharose column (30 x 1.5cm) pre-equilibrated in buffer A (pH 6.5) at 48ml/hour. The column was washed with at least one column volume of buffer A to remove unbound material and the bound protein eluted with a 400ml gradient of 0-0.5M NaCl in buffer A (pH 6.5) at 16-24 ml/hour. SDS PAGE (section 2.9.2) was used to analyse column fractions for the presence of methylase (methylase was easily visible in a crude extract as a result of the overexpression system). Those fractions containing methylase were pooled and dialysed for 4 hours against buffer A (2 l, pH 8.0) to remove NaCl.

4) The dialysed fractions were applied to a heparin-agarose column (12 x 1.5cm) pre-equilibrated in buffer A (pH 8.0) at 48ml/hour. The column was washed and a 400ml gradient of 0-1.0M NaCl in buffer A (pH 8.0) eluted the proteins at 16-24 ml/hour. Fractions containing methylase were pooled and dialysed for 4 hours against buffer A (2 l, pH 8.0) to remove salt.

5) The dialysed fractions were applied to a DEAE-Sepharose column (20 x 1.4cm) pre-equilibrated in buffer A (pH 8.0) at 48ml/hour. The column was washed and a 400ml gradient of 0-0.5M NaCl in buffer A (pH 8.0) eluted the proteins at 16-24 ml/hour. The two peaks containing methylase were pooled separately and

precipitated by adding $(NH_4)_2 SO_4$ to 80%. The pellets obtained after centrifugation were resuspended in buffer A (3-5ml) containing 50% glycerol. Purified protein was stored at -20°C.

2.9.2 Polyacrylamide Gel Electrophoresis (PAGE)

Non-denaturing PAGE was used in gel retardation and DNA methylase assays. Discontinuous sodium dodecyl sulphate (SDS) PAGE was used to analyse protein fractions during purification and also to analyse the products of UV cross linking experiments using methylase and [*methyl*-^{3}H]AdoMet.

A) Non-denaturing PAGE

20xTBE (20ml), acrylamide-bisacrylamide (8.4ml, 30:0.8), H_2O (38.9ml), ammonium persulphate (200µl, 25%), TEMED (15µl)

The volumes above correspond to 5% gels of 2mm thickness, as used for gel retardation. For 5% gels of 1mm thickness, as used for DNA methylase assays, the volumes were halved.

Loading buffer (2x): Tris (0.125M, pH 8.0), glycerol (20%)

Running buffer: 1xTBE

Electrophoresis of both gel retardation and DNA methylase activity assays was monitored by running loading buffer containing bromophenol blue (2.5mg/ml) and xylene cyanol FF (2.5mg/ml) in a spare lane on the gels.

B) SDS PAGE (Laemmli, 1970)

Resolving Gel: Tris-HCl (3.75ml, 3M, pH 8.8), acrylamide-bisacrylamide (10ml, 30:0.8), SDS (0.3ml, 10%), H_2O (14.45ml), ammonium persulphate (90µl, 25%), TEMED (15µl)

Stacking Gel: Tris-HCl (5ml, 0.5M, pH 6.8), acrylamide-bisacrylamide (2.5ml, 30:0.8), SDS (0.2ml, 10%), H_2O (11.3ml), ammonium persulphate (60µl, 25%), TEMED (15µl)

The volumes above correspond to 10% gels of 2mm thickness. For gels of 1mm thickness, as used for the analysis of crosslinked products containing [*methyl*-^{3}H]AdoMet, the volumes were halved.

Loading buffer (2x): Tris-HCl (0.125M, pH 6.8), glycerol (10%), SDS (2%), β-mercaptoethanol (0.7M), bromophenol blue (2.5mg/ml). For fluorography, dye was not included in the loading buffer as it may quench the fluorescence.

Running buffer: Tris (0.025M, pH 8.3), glycine (0.192M), SDS(0.1%)

Coomassie Stain: methanol (10%), acetic acid (10%), coomassie blue stain (1g/l)
Destain: methanol (10%), acetic acid (10%)

1) Samples were mixed with an equal volume of loading buffer and heated in a boiling water bath for 5 minutes and immediately loaded. Gels were run at 30mA.
2) Gels were coomassie stained (overnight) by gentle swirling and excess stain removed by gentle swirling in several changes of destain. Coomassie staining of fluorography gels, if required, was done after having obtained a fluorographic image.

2.9.3 **Desalting of Methylase** (Removal of glycerol)

Protein was prepared for experimentation by desalting on a G-25 Sephadex gel filtration column (PD10 from Pharmacia), the buffer used being dependent on the particular experiment. The concentration and quality of the protein obtained was determined by UV absorption spectrum measurement. A peak-to-trough (A_{280}: A_{250}) ratio of 2.5 - 2.7 was usual. For protein concentration determination (Sober, 1970) the following theoretical extinction coefficients were calculated (at 280nm for a 1mg/ml solution in a 1cm path length):

Methylase	MW M_2S	Molar Absorbance at 280nM M_2S	**A_{280} 1mg/ml M_2S**	MW M_1S	Molar Absorbance at 280nM M_1S	A_{280} 1mg/ml M_1S
wild-type	169914	143969.1	**0.847**	110625	85118.4	0.769
F269Y	169946	146361.7	**0.861**	110641	86314.7	0.780
F269W	169992	155085.7	**0.912**	110664	90676.7	0.819
F269C	169794	143967.7	**0.848**	110565	85117.7	0.770
F269G	169734	143967.7	**0.848**	110535	85117.7	0.770
N266D	169916	143969.1	**0.847**	110626	85118.4	0.769
N266D* F269Y	169948	146361.7	**0.861**	110642	86314.7	0.780
G177D	170030	143969.1	**0.847**	110683	85118.4	0.769

* Not purified

2.9.4.1 Methylase Activity Assay by Overlap of *Hpa* I and M.*Eco* KI Recognition Sites

This assay is dependent upon the overlap between the *Hpa* I and M.*Eco* KI target sites (section 3.5). Restriction by *Hpa* I is inversely proportional to the methylation within the trimeric component of the M.*Eco* KI recognition site.

Oligonucleotides used in the Assay of M.*Eco* KI DNA Methylase Activity (Hpa I linked Assay)

Oligonucleotide	Oligonucleotide Sequence
Hemimethylated substrate	M.*Eco* KI
720/ 721	5'-TGTCTAGATATCGG**GTTAAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCC**CAATTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5' ↑ *Hpa* I CH_3
Unmethylated substrate	
720/ 856	5'-TGTCTAGATATCGG**GTTAAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCC**CAATTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5'

1) Methylase was desalted in the following buffer (buffer M, 1x):
100mM Tris (pH 7.9), NaCl (100mM), $MgCl_2$ (7mM), β-mercaptoethanol (10mM), BSA (0.2mg/ml).

2) Reactions were set up by adding the components in the following order:

Component	Stock Concentration	Volume	Final Concentration
Buffer M	1.9x	8µl	1x
AdoMet	1.1mM	4.5µl	200µM
Methylase	250nM	10µl	100nM
[γ-^{32}P] DNA*	22.5µM	2.5µl	2.25µM

(Reaction volume 25µl)

* Oligonucleotide 720 was end labelled with [γ-^{32}P] as described in section 2.6.13 before hybridisation to either 721 or 856.

3) Reactions were initiated (time 0) by the addition of DNA. At selected time points during the reaction, 2µl aliquots were withdrawn and methylation stopped by incubation in a water bath at 80°C for 15 minutes.

4) After allowing the aliquots to cool to room temperature, the following were added:

1µl NEB buffer 3 (10x stock)

6.5µl H_2O

0.5 µl (5U) of *Hpa* I restriction endonuclease

The restriction digest was incubated at 37°C for 2 hours.

5) Reaction products were separated by polyacrylamide electrophoresis (section 2.9.2) and autoradiographic images obtained at -70°C.

2.9.4.2 Methylase Activity Assay by fluorographic Detection of [*methyl* - ^{3}H] AdoMet

This method is based on the fluorographic detection of [*methyl* - ^{3}H] transfer from [*methyl* - ^{3}H]AdoMet onto a DNA substrate.

Oligonucleotides used in the Assay of DNA Methylase by fluorographic Detection of [*methyl* - ^{3}H] Transfer

Oligonucleotide	Oligonucleotide Sequence
Hemimethylated substrate	
860/ 409	5' -TGTCTAGATATCGGCCTA^{CH_3}**AAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCCGGA**TTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5'
Unmethylated substrate	
408/ 409	5' -TGTCTAGATATCGGCCT**AAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCCGGA**TTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5'

1) Methylase was prepared in the following buffer (buffer M, 1x):
100mM Tris (pH 7.9), NaCl (100mM), $MgCl_2$ (7mM), β-mercaptoethanol (10mM), BSA (0.2mg/ml). Assays were performed at either 37°C or 25°C. Mutant enzymes were assayed in tandem with the wild-type in order to provide a control.

2) Reactions were set up by adding the components in the following order:

Component	Stock Concentration	Volume	Final Concentration
Buffer M	2.2x	3.75µl	1x
[*methyl* - ^{3}H]AdoMet	11.8µM	3µl	2.4µM
Enzyme	222nM	6.75µl	100nM
DNA	22.5µM	1.5µl	2.25µM

(Reaction volume 15µl)

3) Reactions were initiated (time 0) by the addition of DNA. At selected time points during the reaction, 2µl aliquots were withdrawn and the reaction stopped by incubation in a water bath at 80°C for 15 minutes. Aliquots were stored at -70°C until required.

4) Assay products were analysed by fluorographic detection of ^{3}H labelled DNA. Aliquots were run on non-denaturing polyacrylamide gels (section 2.9.2).

5) Gels were fixed (10% methanol, 10% acetic acid) for 30 minutes in order to minimise any diffusion of the products during the subsequent step. Gels were then soaked in Amplify (Amersham) for 30 minutes to enhance the emitted fluorescence.

6) Gels were dried at 65°C for 2 hours and the signal detected by fluorography at -70°C using preflashed Hyperfilm (Amersham). Signals were quantified by densitometry using a Shimadzu dual wavelength thin layer chromatoscanner (model CS-930) densitometer. Multiple exposures were taken in order to obtain signals within the linear range of the instrument. Enzyme activity was determined by plotting intensity vs. time.

2.9.5 Gel Retardation Assay

1) Methylase was prepared in the following buffer: Tris (20mM, pH 8.0), NaCl (100mM), $MgCl_2$ (7mM). Samples for gel retardation contained from 0-100nM methylase, 1nM DNA and 5% glycerol and were incubated for 10 minutes at 22°C prior to electrophoresis. The cofactor AdoMet, if present, was at a concentration of 100µM. In order to measure relative dissociation constants (K_d values), a constant amount of DNA was titrated with increasing amounts of methylase.

2) Free DNA and bound DNA were separated on 5% non-denaturing polyacrylamide gels (section 2.9 2). Gels were run at 30 mA for approximately 2 hours, dried at 80°C for 1.5 hours and autoradiographed using preflashed film.

3) K_d values for binding were estimated by determination of the methylase concentration needed to complex half of the DNA under conditions where the methylase was in excess over the DNA (Hendrickson and Schleif, 1984). K_d values were determined from plots of % free DNA vs. methylase concentration. Free DNA was quantified by densitometry (section 2.9.4). Depletion of free DNA provides a more accurate measure of complex formation than bound DNA because free DNA cannot change once it has entered the gel. Bound DNA, however, can dissociate within the gel resulting in a smearing effect below the complex making quantification more difficult (Fried, 1989).

2.9.6 AdoMet Binding

2.9.6.1 AdoMet Binding by Fluorescence of the extrinsic Fluorophore ANS

1) Methylase was prepared in buffer A, 200mM NaCl, pH 6.5 (section 2.9.1). AdoMet (1μM) and ANS (10mM) were made up in the same buffer and their concentrations determined by spectroscopy (for AdoMet, A_{260} of a 1M solution = 15400: for ANS, A_{350} of a 1M solution = 4950; Sober, 1970).

2) A 400μl solution containing 1μM methylase and 50μM ANS was titrated with AdoMet over the concentration range 0-300μM in a 400μl quartz microcuvette. Excitation was at 395nm, emission was at 480nm, bandwidths were 10nm. Fluorescence spectra and titrations were measured using a 2mm pathlength with a Perkin Elmer LS50 fluorimeter. Experiments were performed at 25°C and temperature was controlled to ±0.1°C.

3) Data were corrected for the fluorescence emission of the buffer and unbound ANS.

4) Fluorescence data were analysed using the Grafit data analysis programme (Leatherbarrow, 1990) with a ligand-binding equation for identical non-interacting sites.

2.9.6.2 AdoMet Binding by Gel Filtration

1) Methylase (50μl, 6μM) was added to [*methyl* - ^{3}H]AdoMet (1μl, 12μM) and samples applied to a Sephadex G25 column (5mm x 14.6cm) equilibrated in buffer A, 200mM NaCl, pH 6.5 (section 2.9.1).

2) Fractions of two drops (approximately 100μl) were collected and after adding 2.5ml scintillation fluid, the ^{3}H was counted in a Beckmann scintillation counter (model LS7000).

3) Counts per minute were plotted vs. fraction number in order to compare mutant enzymes with the wild-type.

2.9.6.3 AdoMet Binding by UV Crosslinking of Protein to [*methyl* - ^{3}H]AdoMet

1) Methylase was prepared in buffer M (section 2.9.4). Methylase (600nM) and [*methyl* - ^{3}H]AdoMet (2.4μM) were irradiated on ice under UV light (254nm) in a Stratagene X-linker for 10 minutes. Samples were placed as close as possible to the UV light source.

2) The irradiated samples were separated on SDS PAGE (section 2.9.2). Gels were

fixed (30 minutes), soaked in amplify (30 minutes) and a fluorographic image obtained as described in section 2.9.4. Signals were quantified by densitometry using a Shimadzu dual wavelength thin layer chromatoscanner (model CS-930) densitometer. Multiple fluorographic exposures were taken in order to obtain signals within the linear range of the instrument.

2.9.7 Protein Denaturation using Guanidine Hydrochloride

1) Methylase was prepared in buffer A, 200mM NaCl, pH 8.0 (section 2.9). Guanidine hydrochloride (GuCl) was prepared in the same buffer and the concentration determined by refractometry using the following equation (Pace *et al.*, 1989):

$$\text{Molarity} = 57.147(\Delta N) + 38.68(\Delta N^2) - 91.60(\Delta N^3)$$

(where ΔN is the difference in refractive index between the denaturant solution and the buffer at the sodium D line.)

2) Samples of protein (200nM) were incubated overnight with GuCl in order to allow equilibrium to be reached (Pace, 1986). Denaturation experiments with the wild-type enzyme have indicated aggregation when using a protein concentration of 1.8μM but this problem was alleviated upon using a concentration of 200nM (Dryden *et al.*, 1993). Approximately 25 samples were set up in the GuCl concentration range 0-4M in a reaction volume of 400μl.

3) Fluorescence and anisotropy measurements were determined at 20°C (±0.1°C) on a Perkin-Elmer LS50 fluorimeter. Excitation was at 295nm (5nm bandwidth) in order to selectively excite tryptophan residues. Fluorescence emission was measured at 320nm and 370nm (5nm bandwidths) in order to monitor the progress of the red shift occurring upon denaturation. Anisotropy was measured using an excitation wavelength of 295nm (15nm bandwidth) and an emission wavelength of 345nm (20nm bandwidth) in order to follow changes in polarisation throughout the denaturation. Data were corrected for any fluorescence due to the buffer or the GuCl. Denaturation curves were compared to those of the wild-type in order to indicate any significant differences.

Chapter Three

RESULTS

3.1 Introduction

The following experiments address the roles played in the *Eco* KI methylase by two conserved motifs common to N6-adenine methylases. A combination of *in vivo* and *in vitro* techniques have been used to investigate the effect upon methylase activity, conformational stability, AdoMet binding and DNA binding of site-directed mutations made within these two motifs.

The first of these motifs (region I) is common not only to N6-adenine methylases but to all methylases that use AdoMet as the methyl donor (Ingrosso *et al.*, 1989). The universal occurrence of this motif in such enzymes suggests a key role in the interaction with AdoMet.

The second of these motifs (region II) is found in both N6-adenine methylases and also in N4-cytosine methylases, both of which catalyse the methylation of an extracyclic amino group. The motif displays a low degree of similarity to the characterised P C motif found in C5-cytosine methylases and therefore may fulfil a similar function, i.e. a central role in catalysis (Klimasauskas *et al.*, 1989).

3.2 Generation of Site-Directed Mutations within *hsdM*

The 1.4kb *Sma* I-*Bam* HI containing the major portion of *Eco* KI *hsdM* was cloned in M13mp19 and used as substrate for site-directed mutagenesis (figure 3.1). Table 3.1 details the oligonucleotides used to create these mutations within *hsdM*. Figure 3.2 details the nucleotide changes as observed by DNA sequencing.

One mutation, G177D, was generated within region I, (the mutagenic oligonucleotide 709J was designed to generate two mutations at this position but only one of these was found). This Gly residue is very highly conserved in AdoMet dependent methyltransferases (table 1.3) and substitution of the larger acidic residue Asp for Gly might be anticipated to have some effect upon enzyme activity.

Six mutations were generated within the region II motif, Asn Pro Pro Phe (amino acids 266-269). Three of these mutations remained within the overall consensus found for this motif; the single mutants N266D and F269Y, and the double mutant N266D/F269Y. These mutations go some way to defining the degree of freedom allowed within the overall consensus. The other three mutants made in region II were outside the consensus; F269C, F269G and F269W. Substitution of the smaller residues Cys and Gly for the relatively bulky amino acid Phe might be anticipated to have some effect upon enzyme activity. Substitution of Trp for Phe represents a more conserved change and might be anticipated to have less of an effect than substitution of Cys or Gly.

Figure 3.1

The *hsd* Region of *E.coli* K12 and the Cloning Steps used in this Work.

Figure 3.1

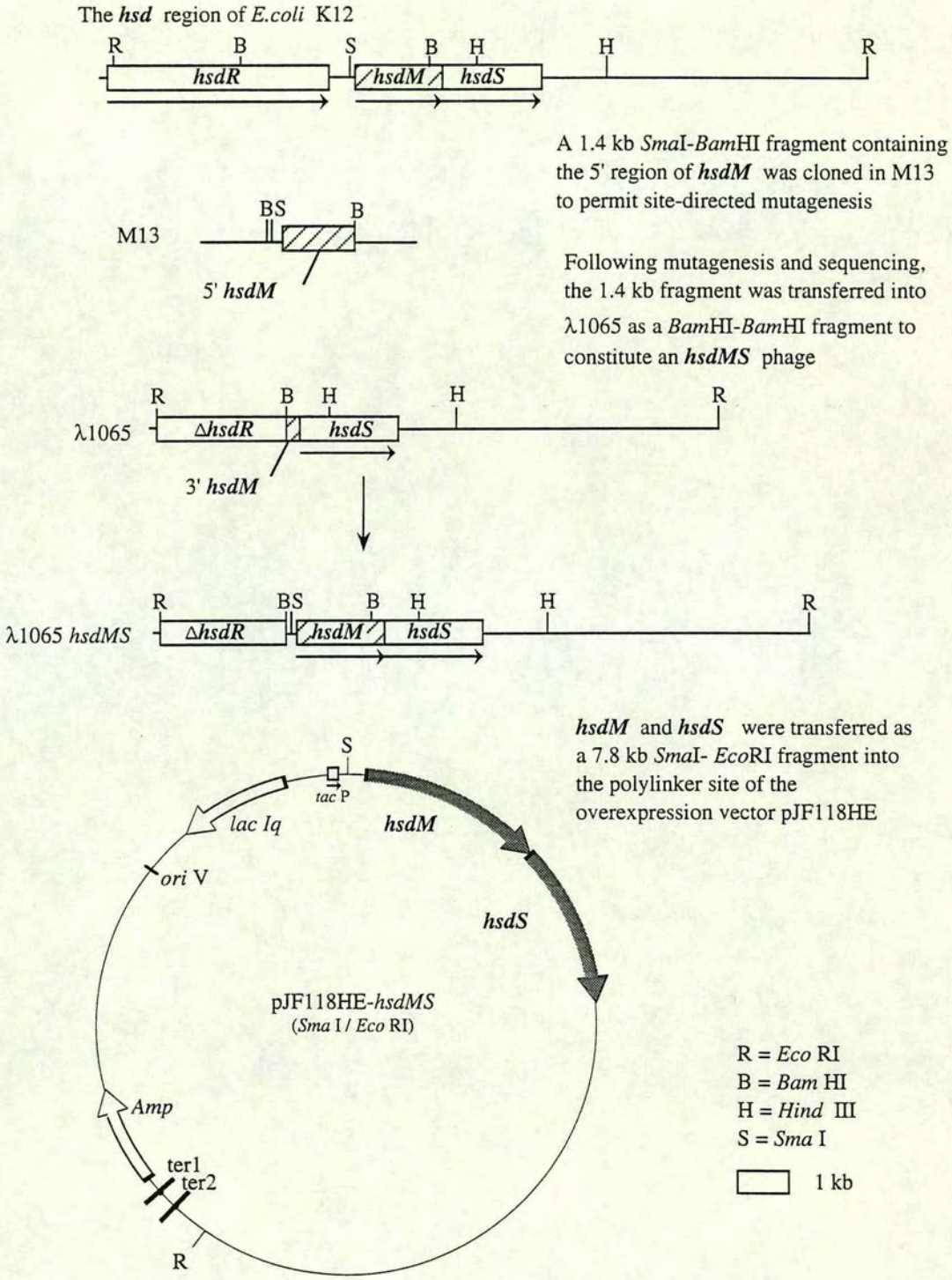

The *hsd* region of *E.coli* K12
R
B
S
B
H
H
R
hsdR
hsdM
hsdS
A 1.4 kb *Sma*I-*Bam*HI fragment containing the 5' region of ***hsdM*** was cloned in M13 to permit site-directed mutagenesis
M13
BS
B
5' *hsdM*
Following mutagenesis and sequencing, the 1.4 kb fragment was transferred into λ1065 as a *Bam*HI-*Bam*HI fragment to constitute an ***hsdMS*** phage
λ1065
ΔhsdR
hsdS
3' *hsdM*
λ1065 *hsdMS*
hsdM and ***hsdS*** were transferred as a 7.8 kb *Sma*I- *Eco*RI fragment into the polylinker site of the overexpression vector pJF118HE
tac P
lac Iq
ori V
Amp
ter1
ter2
pJF118HE-*hsdMS*
(*Sma* I / *Eco* RI)
R = *Eco* RI
B = *Bam* HI
H = *Hind* III
S = *Sma* I
1 kb

Table 3.1 Site-directed Mutations made within *hsdM*

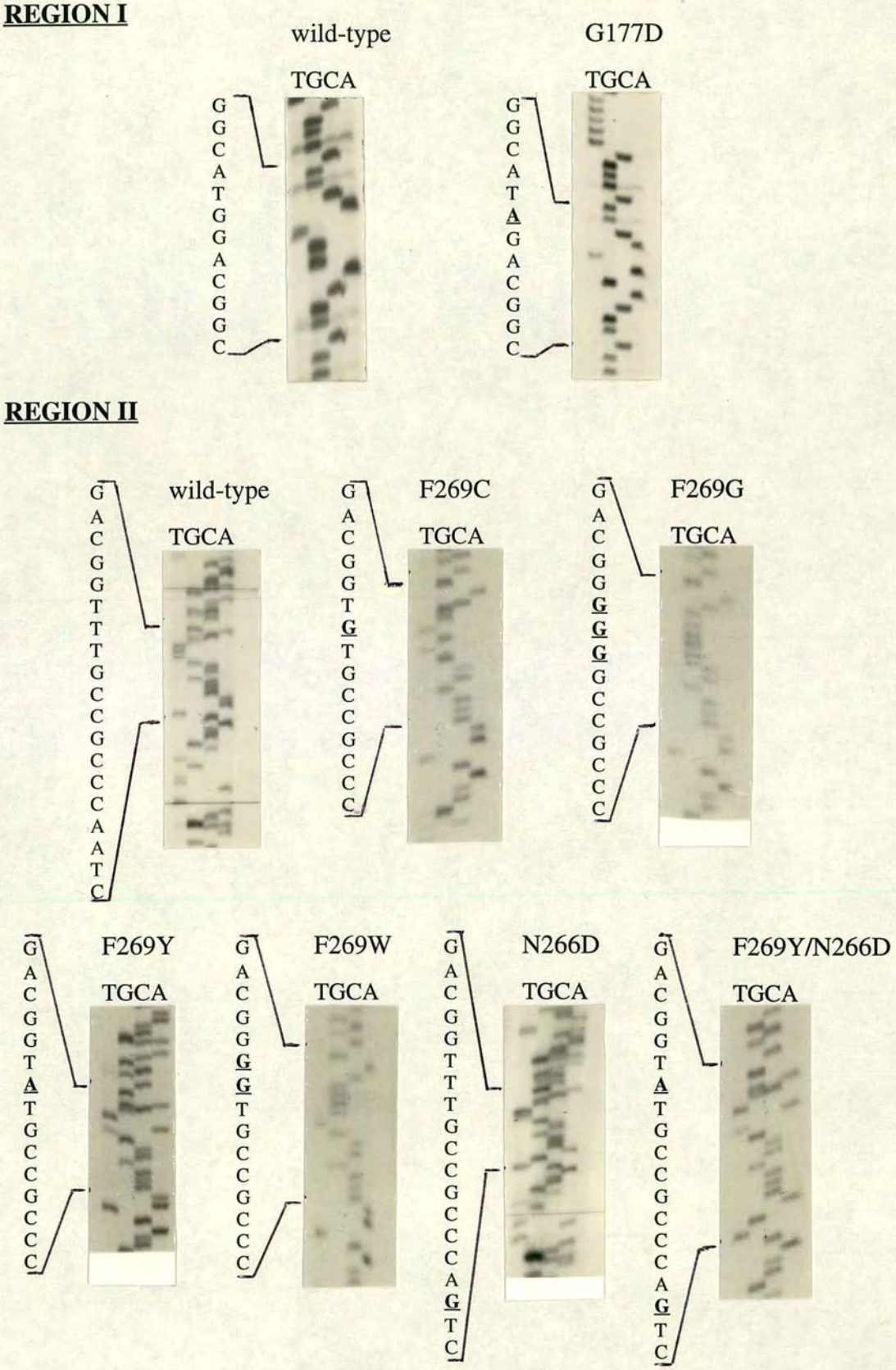
	Oligonucleotide Sequence (Amino Acid Sequence)		
REGION I			
Wild-type	5'-GACCCGGCGGCA**GGT**ACGGCGG AspProAlaAla**Gly**ThrAla **177**		
Oligonucleotide			
709J	**C** 5'-CGGCGGCA**G** **T**ACGGCGG **A**		
	Ala* (GCT)	→	G177A*
	AlaAla Thr		
	Asp (GAT)	→	G177D
REGION II			
Wild-type	5'-CGCCACT**AAC**CCGCCG**TTT**GGCAGCGCCGCAG **Asn**ProPro**Phe** 266 269		
Oligonucleotide			
252E	**G** **G** 5'-CTAACCCGCCG **G** GGCAGCGCCGCAG **T** **T**		
	Gly (GGG)	→	F269G
	AsnProPro**Trp** (TGG)	→	F269W
	Cys (TGT)	→	F269C
228B	5'-CCGCCG**TAT**GGCAGC		
	ProPro**Tyr**	→	F269Y
229B	5'-CGCCACT**GAC**CCGCC		
	AspPro	→	N266D
227B	5'-CGCCACT**GAC**CCGCCG**TAT**GGCAGC		
	AspProPro**Tyr**	→	N266D/F269Y

* No examples of this mutation (G177A) were found despite the degeneracy of the oligonucleotide sequence at this position.

Figure 3.2

DNA Sequencing of the site-directed Mutations made in Regions I and II.

Figure 3.2


REGION I
wild-type
TGCA
G177D
TGCA
REGION II
wild-type
TGCA
F269C
TGCA
F269G
TGCA
F269Y
TGCA
F269W
TGCA
N266D
TGCA
F269Y/N266D
TGCA

3.3 *In vivo* Methylase Activity

A λ phage encoding functional *hsdMS* genes will methylate its DNA during replication and will therefore be protected from host-encoded restriction of the same specificity. Conversely, a phage encoding non-functional *hsdMS* genes will not self methylate and its DNA will be restricted by such a host. Comparison of the titre of phage encoding *hsdMS* genes on restricting and non-restricting hosts of the same specificity provides an *in vivo* measure of methylase activity.

After isolating the desired mutations in *hsdM* the 1.4kb *Sma* I-*Eco* RI fragments cloned in M13mp19 were transferred as 1.4kb *Bam* HI-*Bam* HI fragments into the λ vector, NM1065 (figure 3.1).

Following transfer into the λ vector the *Bam* HI fragments were recloned in M13mp19 and sequenced in entirety to ensure that no unwanted mutations were present. None was found.

The vector λNM1065 contains the remainder of *hsdM* and *hsdS* (figure 3.1). The mutations in *hsdM* encoded by the NM1065 derivatives were used to provide an *in vivo* determination of their methylase activity. The bacterial strain C600 is K-restricting whereas the strain NM679 is deleted for the whole of the *hsdK* region. Plating efficiencies (e.o.p) were determined for each λ*hsdMS* mutant derivative (table 3.2). The *in vivo* methylation results indicate three classes of mutant. The first class are those mutants which have the same methylation activity as the wild-type methylase, that is, they have an equal e.o.p. on C600 as the wild-type. This class comprises only F269Y. The second class are those mutants which have no methylase activity, that is, they have the same e.o.p. on C600 as does the vector. This class comprises G177D, N266D, N266D/F269Y, F269C and F269G. The third class are those mutants which have an intermediate methylation activity and this class comprises only F269W.

These results indicate that mutations in region II which remain within the overall consensus do not necessarily retain any *in vivo* methylation activity. Of these mutations only F269Y is active. This would suggest that the degree of freedom within the consensus is limited. Of the mutations within region II which are outside the consensus, F269C and F269G show no activity whereas F269W shows

partial activity. This may indicate that the size of the residue at this position is of critical importance.

The mutation in region I, G177D, shows no *in vivo* activity confirming the expectation that replacement of this very highly conserved residue for a larger charged residue would have an effect upon enzyme activity.

Table 3.2 Efficiencies of Plating for λ NM1065 Derivatives

Phage construct[a]	Titre on C600	Titre on 679	Efficiency of Plating
λNM1065	$1x10^{6}$	$1x10^{10}$	$1x10^{-4}$
λNM1065-wt	$4x10^{9}$	$2x10^{10}$	$2x10^{-1}$
λNM1065-G177D	$1.4x10^{6}$	$1x10^{10}$	$1.4x10^{-4}$
λNM1065-N266D	$3.6x10^{6}$	$3x10^{10}$	$1.2x10^{-4}$
λNM1065-F269C	$6x10^{6}$	$5x10^{10}$	$1.2x10^{-4}$
λNM1065-F269G	$4.5x10^{6}$	$3x10^{10}$	$1.5x10^{-4}$
λNM1065-F269W	$1.1x10^{8}$	$1x10^{10}$	$1.1x10^{-2}$
λNM1065-F269Y	$8.8x10^{9}$	$2x10^{10}$	$4.4x10^{-1}$
λNM1065-N266D/F269Y	$7.8x10^{6}$	$6x10^{10}$	$1.3x10^{-4}$

[a] All phage had previously been grown on the *hsdK*$^{-}$ host, NM679.

3.4 Overexpression and Purification of Mutant Methylases

All the mutations (except N266D/F269Y) were transferred from λ NM1065 within the 7.8kb *Sma* I-*Eco* RI fragment encoding the HsdM and S polypeptides to the over expression plasmid pJF118HE (Furst *et al.*, 1986) in order to examine their *in vitro* properties (figure 3.1).

All but one of the mutant enzymes were found in the soluble fraction when expressed at 37°C. The region I mutant G177D was found predominantly in the insoluble fraction when expressed at this temperature. However when expressed at 25°C a substantial proportion (more than half) was located in the soluble fraction thus allowing its purification using the same steps used for the other mutants.

In each case the mutant methylases were purified in the same manner, eluting from each of the three columns at characteristic NaCl concentrations (figure 3.3). Protein eluted from the first column, CM-Sepharose, as a broad peak at a salt concentration of between 0.1 and 0.2M NaCl. Gradient elution from the second column, heparin-agarose, resulted in two protein peaks eluting at NaCl concentrations of approximately 0.23-0.27M and 0.27-0.35M. The two peaks were pooled and applied to the third column, DEAE-Sepharose. Gradient elution again resulted in two protein peaks. The peaks eluted at NaCl concentrations of approximately 0.23-0.3M and 0.3-0.35M. SDS PAGE indicated the second of the two peaks to be predominantly of the form M_2S_1. The fractions from these two peaks were pooled separately and experiments confined to the M_2S_1 species since this has been shown to be the active species in the wild-type enzyme (Dryden *et al.*, 1993).

All but one of the mutant enzymes therefore appear to be very similar to the wild-type enzyme in their manner of purification, indicative of proteins which have no major structural differences. This would suggest that the *in vivo* inactivity of the F269G, F269C, F269W and N266D mutants is the result of a local alteration, not affecting the folding of subunits or the final assembly of HsdM and HsdS.

The temperature sensitivity of G177D, however, may suggest a radical alteration at 37°C to the stability of this mutant protein. Fortunately the majority of

the protein is soluble at 25°C thus allowing recovery and further purification of sufficient quantities as required for an *in vitro* analysis.

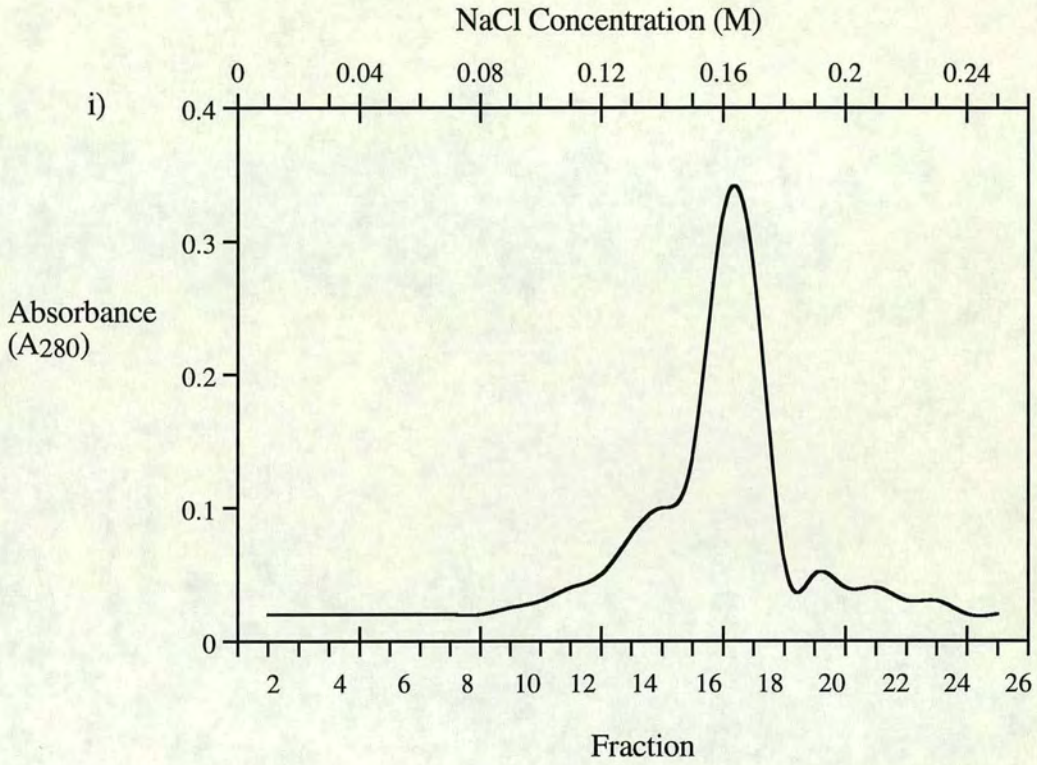
Figure 3.3 (1 of 3)

Methylase Purification: NaCl Elution of Methylase from Column 1, CM-Sepharose.

i) Absorbance at 280nm vs. fraction number and NaCl concentration.

ii) SDS PAGE of eluted fractions.

Figure 3.3 (1 of 3)



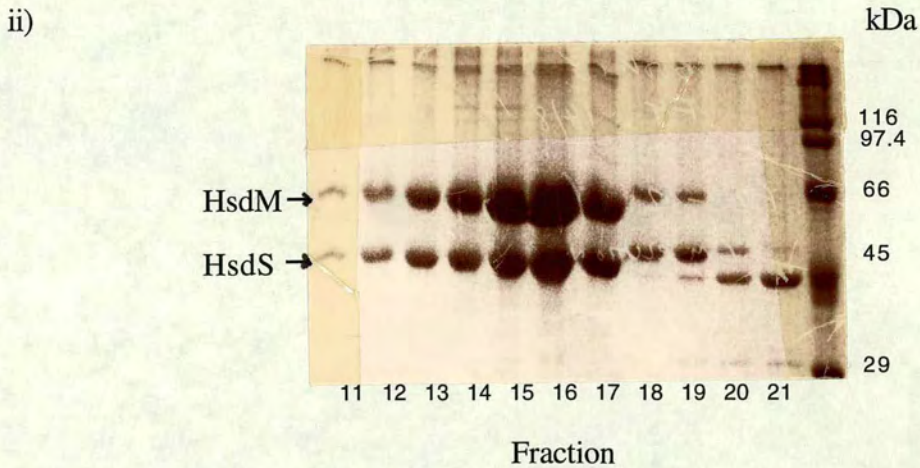


Figure 3.3 (2 of 3)

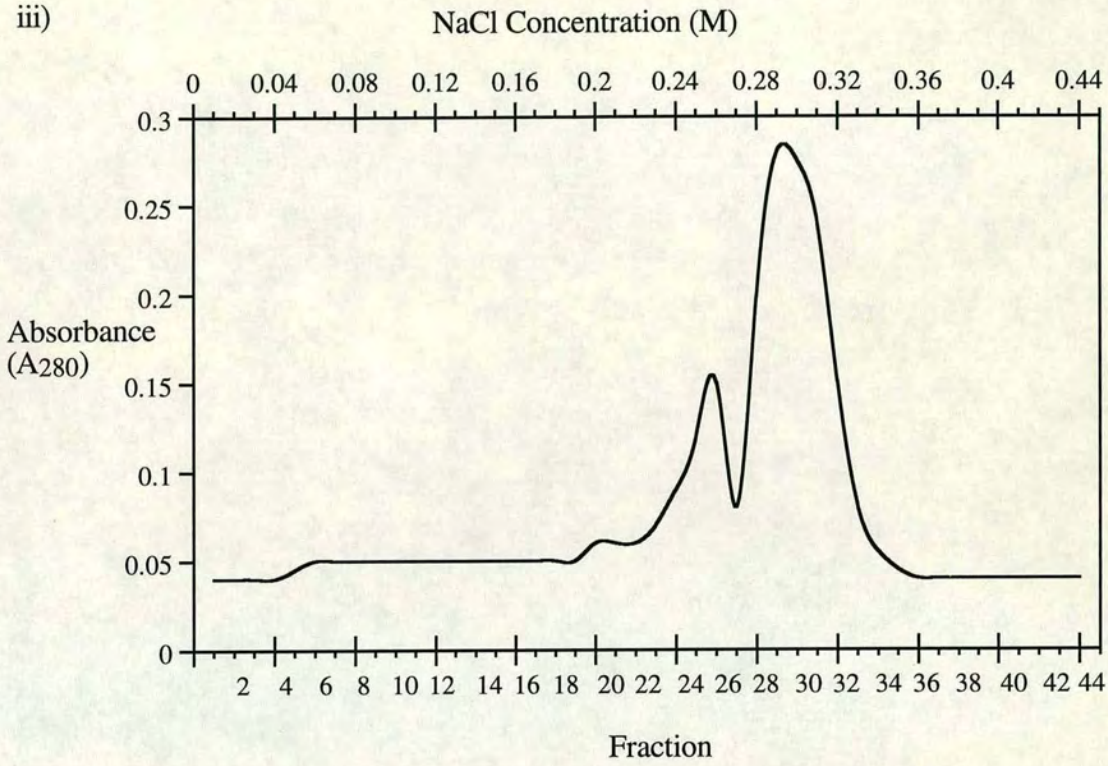
Methylase Purification: NaCl Elution of Methylase from Column 2, Heparin Agarose.

iii) Absorbance at 280nm vs. fraction number and NaCl concentration.

iv) SDS PAGE of eluted fractions.

Figure 3.3 (2 of 3)

iii)



iv)

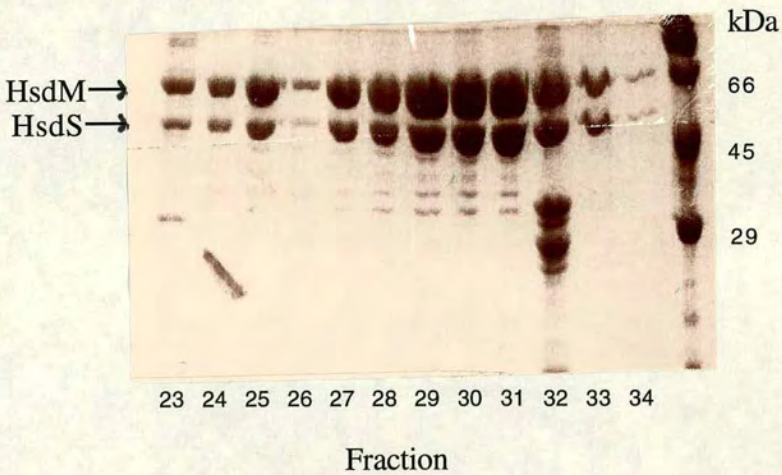


Figure 3.3 (3 of 3)

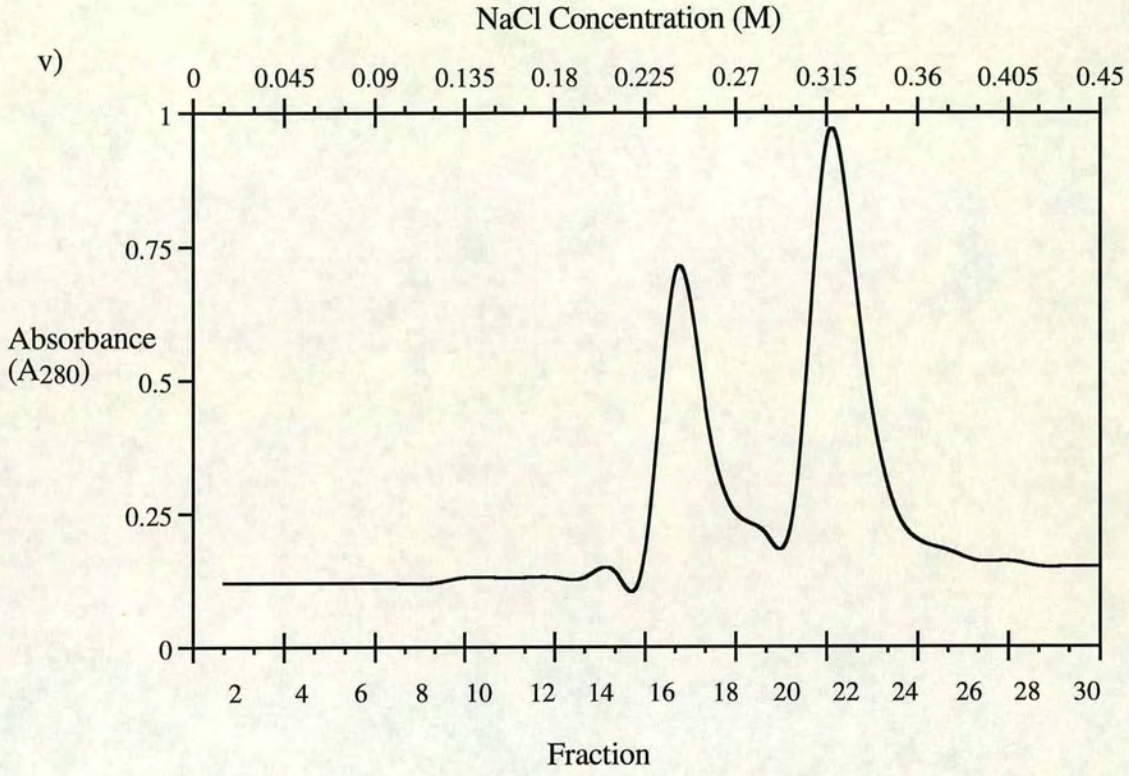
Methylase Purification: NaCl Elution of Methylase from Column 3, DEAE-Sepharose.

v) Absorbance at 280nm vs. fraction number and NaCl concentration.

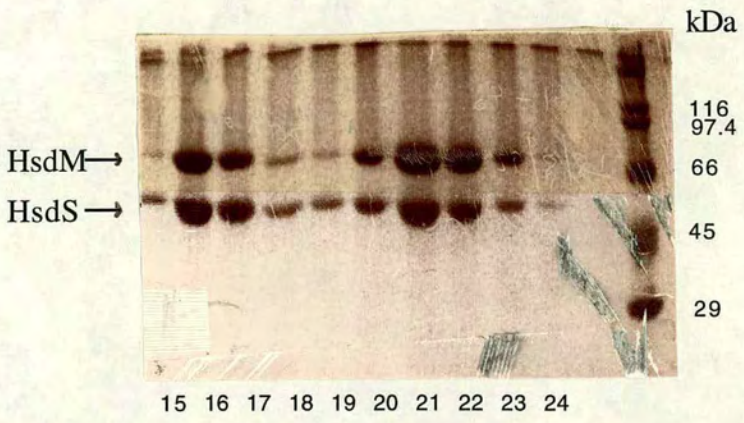
vi) SDS PAGE of eluted fractions.

Figure 3.3 (3 of 3)

v)



vi)



3.5 *In Vitro* Methylase Activity

The IA and IC families of type I R-M systems are characterised by their preference in the methylation reaction for a hemimethylated target site (Suri *et al.*, 1984; Dryden *et al.*, 1993; Taylor *et al.*, 1993). In order to investigate the *in vitro* activity of the mutant enzymes with reference to the wild-type activity and also their preference, if any, for a hemimethylated substrate, an *in vitro* methylase assay was developed. For a full analysis of enzyme activities and substrate preferences, a continuous assay would be of great advantage due to the inherent greater accuracy of such a system over discontinuous systems. So far there are no continuous methods for the assay of DNA methylases. One discontinuous method of assay, originally developed by Rubrin and Modrich (1977), involves scintillation counting of tritiated DNA substrate immobilised on ion-exchange paper. However this method was not used for the assay of M.*Eco* KI as it does not appear to be sensitive enough to detect activity at low enzyme concentrations (Dryden, D.T.F., personal communication).

In this study two further discontinuous assay techniques were devised although only one was used to assay the mutant enzymes. The first of these assays takes advantage of the partial overlap between the target recognition sites of M.*Eco* KI and the type II restriction endonuclease *Hpa* I.

Enzyme	Target Recognition Site
Hpa I	5'-GTTAAC
M.*Eco* KI	5'-AAC(N_6)GTGC

Methylation at the N6 position of the underlined adenine within the *Hpa* I recognition site renders the target resistant to cleavage. This adenine residue, when part of the M.*Eco* KI recognition target, is one of the two target sites for methylation by M.*Eco* KI (the other is on the complementary strand within the tetrameric component). Hence when the target recognition sites overlap as indicated above, restriction by *Hpa* I becomes a function of methylation by M.*Eco* KI. Another type II restriction endonuclease, *Hinc* II, can be used in place of *Hpa* I in this assay. *Hinc* II is a partial isoschizomer of *Hpa* I, having the degenerate recognition sequence, 5'-GTPyPuAC.

In order to assay M.*Eco* KI three oligonucleotides were designed to provide both a hemimethylated (720/721) and an unmethylated (720/856) target. The resulting oligonucleotide duplexes (table 3.3) incorporate the *Hpa* I and M.*Eco* KI target sites in the arrangement indicated above. Both duplexes are 45 bp long and are identical except for the N6-methyladenine within the tetrameric component of the M.*Eco* KI target sequence contained by 720/721. Cleavage of either DNA substrate by *Hpa* I (indicated by ↑) results in the formation of two shorter duplexes (blunt ended) of 17 and 28 bp that can be resolved by PAGE. The extent of restriction of either substrate by *Hpa* I is dictated by the level of methylation by M.*Eco* KI within the trimeric component. This novel method is described in chapter 2 and an example of the method as used to assay the wild-type is indicated in figure 3.4. Figure 3.4 shows that restriction by *Hpa* I of the hemimethylated DNA substrate (720/721) is not detectable after 3 minutes of methylation by M.*Eco* KI. However restriction of the unmethylated substrate (720/856) is evident even after 5 hours indicating that this latter DNA is methylated very slowly by M.*Eco* KI. However the reproducibility of this assay was not good, perhaps due to fluctuations in the specific activity between different batches of *Hpa* I.

A second discontinuous assay was devised based on the fluorographic detection of [*methyl* - ^{3}H] transfer from [*methyl* - ^{3}H]AdoMet onto a DNA substrate. This third method, compared to the two methods described above, generally provided better sensitivity and reproducibility and for these reasons was used to assay the mutant enzymes. Methylase activity was measured on both unmethylated (408/409) and hemimethylated (860/409) oligonucleotide substrates (table 3.4). Both substrates are identical except for the N6-methyladenine within the trimeric component of the M.*Eco* KI target sequence contained by 860/409.

All mutant enzymes were assayed at 37°C except G177D which was assayed at both 37°C and 25°C. G177D is insoluble when purified at 37°C (section 3.4) and for this reason alone it is unlikely to be active at this temperature. However its solubility when purified at 25°C suggests that the extent of activity when assayed at this lower temperature might depend upon the effect of the mutation but not on the insolubility which is associated with the higher temperature.

Methylase activity of the mutants was measured relative to the wild-type activity which was performed in tandem in order to provide a reference (figure 3.5).

For those mutants which demonstrated *in vitro* activity a plot of activity vs. time is shown in figure 3.6. The initial rate of activity of the mutant enzymes was measured from these plots and is expressed as the percentage of wild-type activity in table 3.5.

Table 3.3 **Oligonucleotides used in the Assay of M.*Eco* KI DNA Methylase Activity (*Hpa* I linked Assay)**

Oligonucleotide	Oligonucleotide Sequence
Hemimethylated substrate	
720/ 721	5' -TGTCTAGATATCGG**GTTAAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCC**CAATTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5' M.*Eco* KI; *Hpa* I; CH_3
Unmethylated substrate	
720/ 856	5' -TGTCTAGATATCGG**GTTAAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCC**CAATTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5'

Table 3.4 **Oligonucleotides used in the Assay of M.*Eco* KI DNA Methylase Activity (fluorographic Detection of [*methyl* - ^{3}H] Transfer)**

Oligonucleotide	Oligonucleotide Sequence
Hemimethylated substrate	
860/ 409	5' -TGTCTAGATATCGGCCT**AAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCCGGA**TTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5' CH_3
Unmethylated substrate	
408/ 409	5' -TGTCTAGATATCGGCCT**AAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCCGGA**TTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5'

Figure 3.4

***In vitro* Assay of Wild-Type M.*Eco* KI using the *Hpa* I linked Method.** Restriction by *Hpa* I of the hemimethylated DNA substrate (720/721) is not detectable after 3 minutes of methylation by M.*Eco* KI. However restriction of the unmethylated substrate (720/856) is evident even after 5 hours indicating that this latter DNA is methylated very slowly by M.*Eco* KI.

Figure 3.4

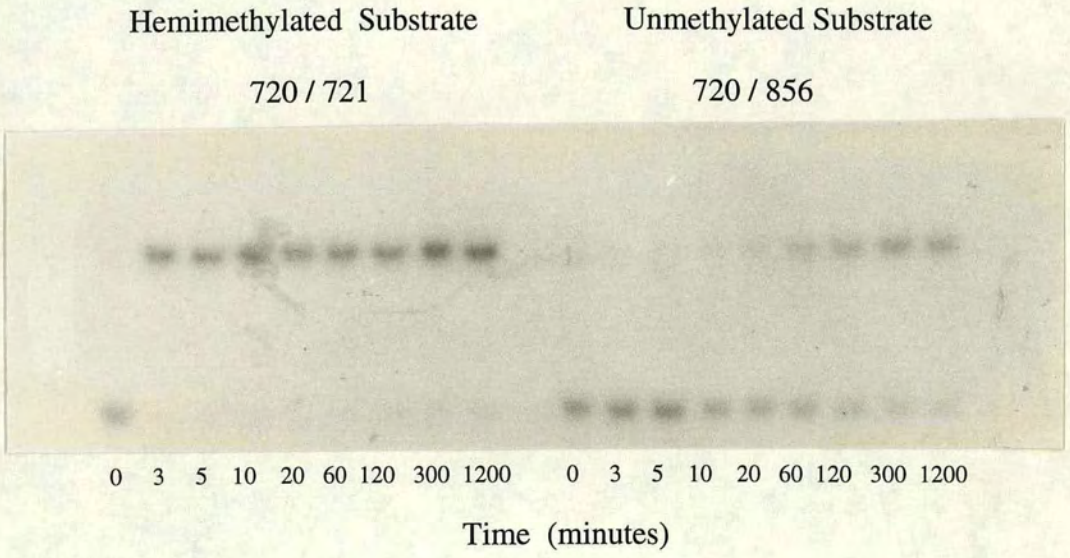

Hemimethylated Substrate
720 / 721
Unmethylated Substrate
720 / 856
0 3 5 10 20 60 120 300 1200
0 3 5 10 20 60 120 300 1200
Time (minutes)

Figure 3.5 (1 of 3)

***In vitro* Assay of mutant Methylases by fluorographic Detection of [*methyl* - ^{3}H] Transfer.** Methylases F269C, F269G, N266D and G177D have no detectable activity on either a hemimethylated (860/409) or an unmethylated (408/409) DNA substrate. Methylases F269Y has activity on both substrates. F269W has activity on a hemimethylated substrate and may have activity on an unmethylated substrate.

I F269C
II F269G

Figure 3.5 (1 of 3)

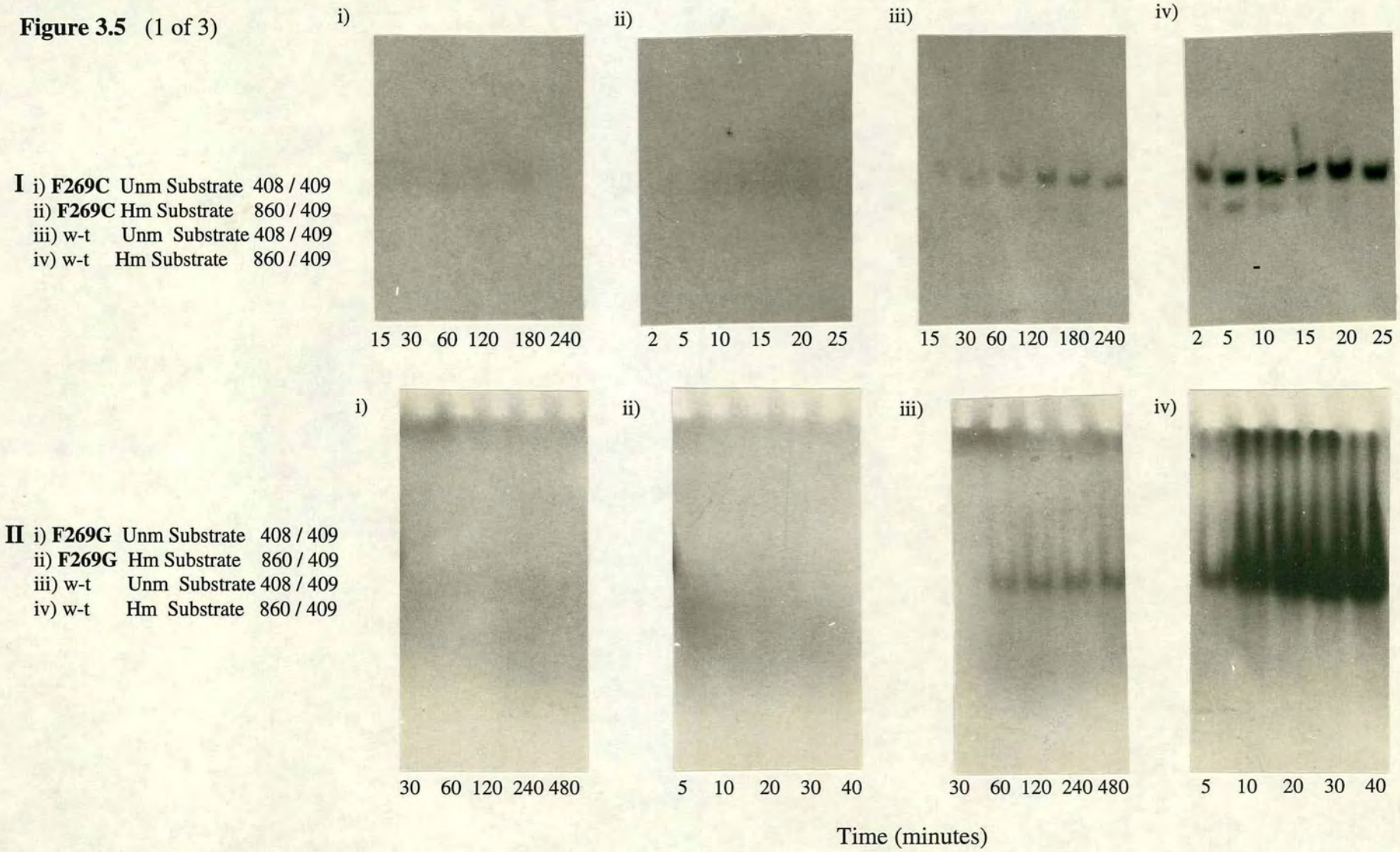
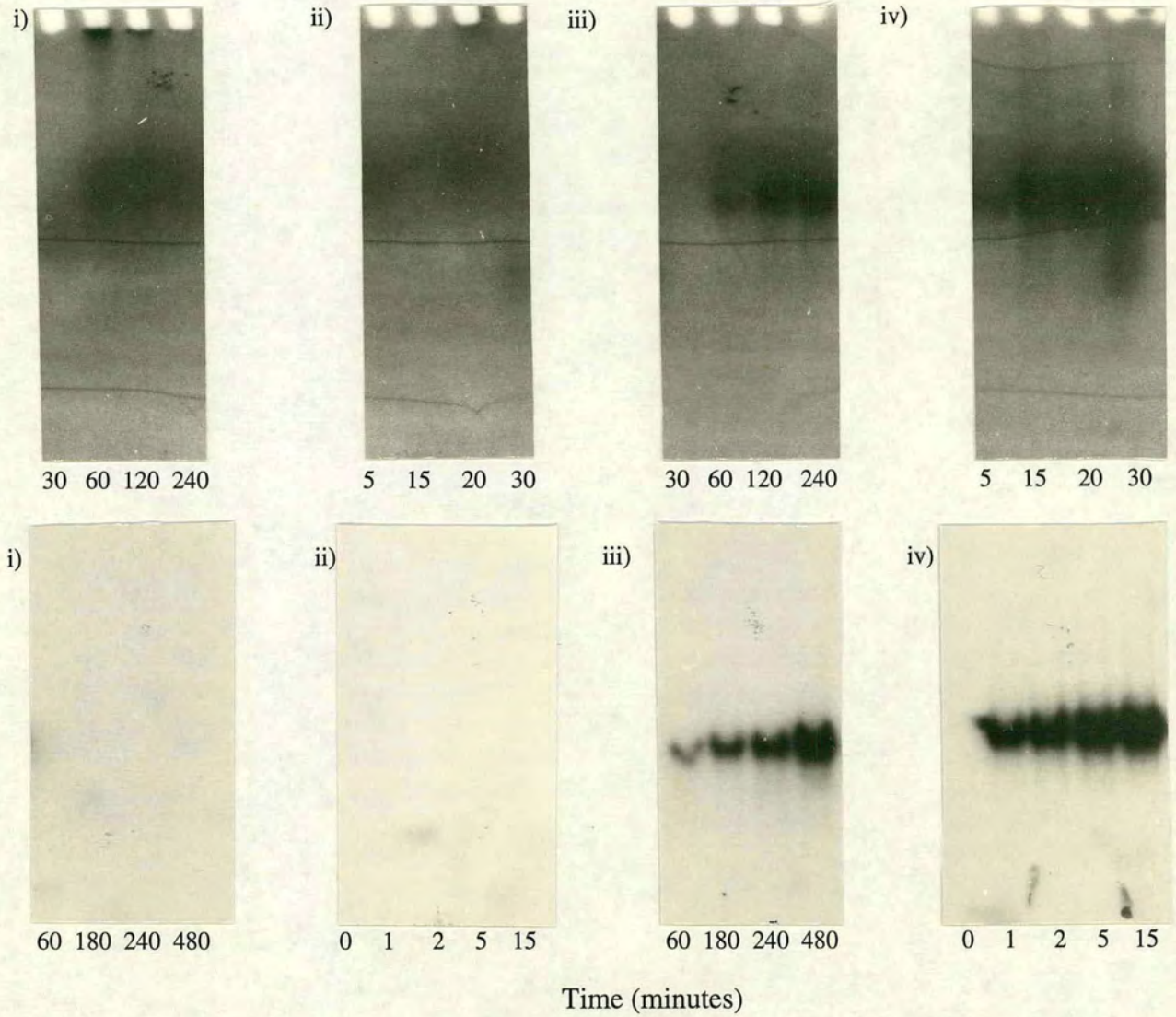


Figure 3.5 (2 of 3)

In vitro **Assay of mutant Methylases by fluorographic Detection of [*methyl* - ^{3}H] Transfer.**

III N266D
IV G177D

Figure 3.5 (2 of 3)



III i) **N266D** Unm Substrate 408 / 409
ii) **N266D** Hm Substrate 860 / 409
iii) w-t Unm Substrate 408 / 409
iv) w-t Hm Substrate 860 / 409

IV i) **G177D** Unm Substrate 408 / 409
ii) **G177D** Hm Substrate 860 / 409
iii) w-t Unm Substrate 408 / 409
iv) w-t Hm Substrate 860 / 409

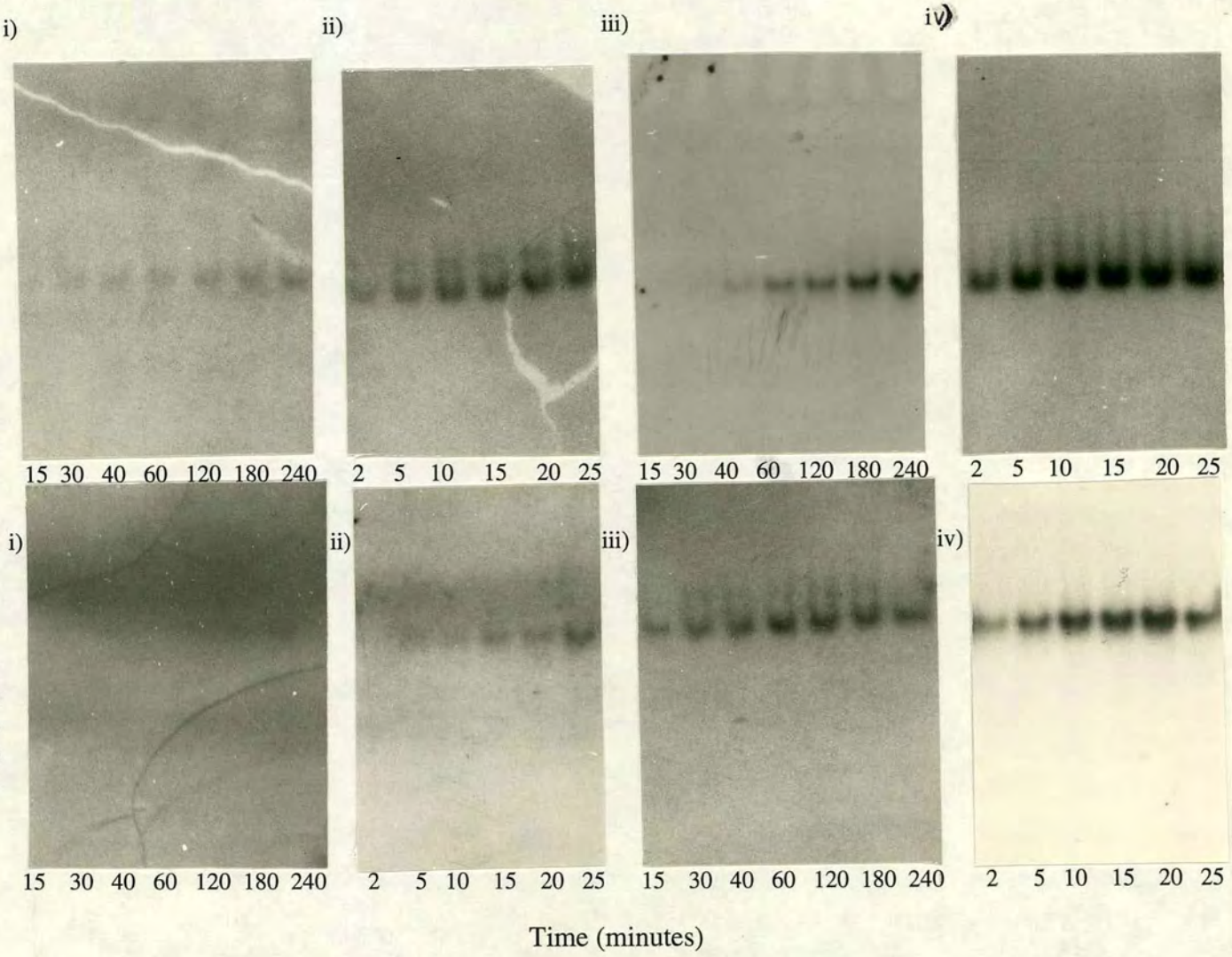
Figure 3.5 (3 of 3)

In vitro **Assay of mutant Methylases by fluorographic Detection of [*methyl* - ^{3}H] Transfer.**

V F269Y

VI F269W

Figure 3.5 (3 of 3)



V i) **F269Y** Unm Substrate 408 / 409
ii) **F269Y** Hm Substrate 860 / 409
iii) w-t Unm Substrate 408 / 409
iv) w- t Hm Substrate 860 / 409

VI i) **F269W** Unm Substrate 408 / 409
ii) **F269W** Hm Substrate 860 / 409
iii) w-t Unm Substrate 408 / 409
iv) w-t Hm Substrate 860 / 409

Figure 3.6 (1 of 3)

In vitro **Methylase Activity (fluorographic Detection of [*methyl* - ^{3}H] Transfer).**

i) F269Y on unmethylated DNA Substrate 408/409.

Figure 3.6 (1 of 3)

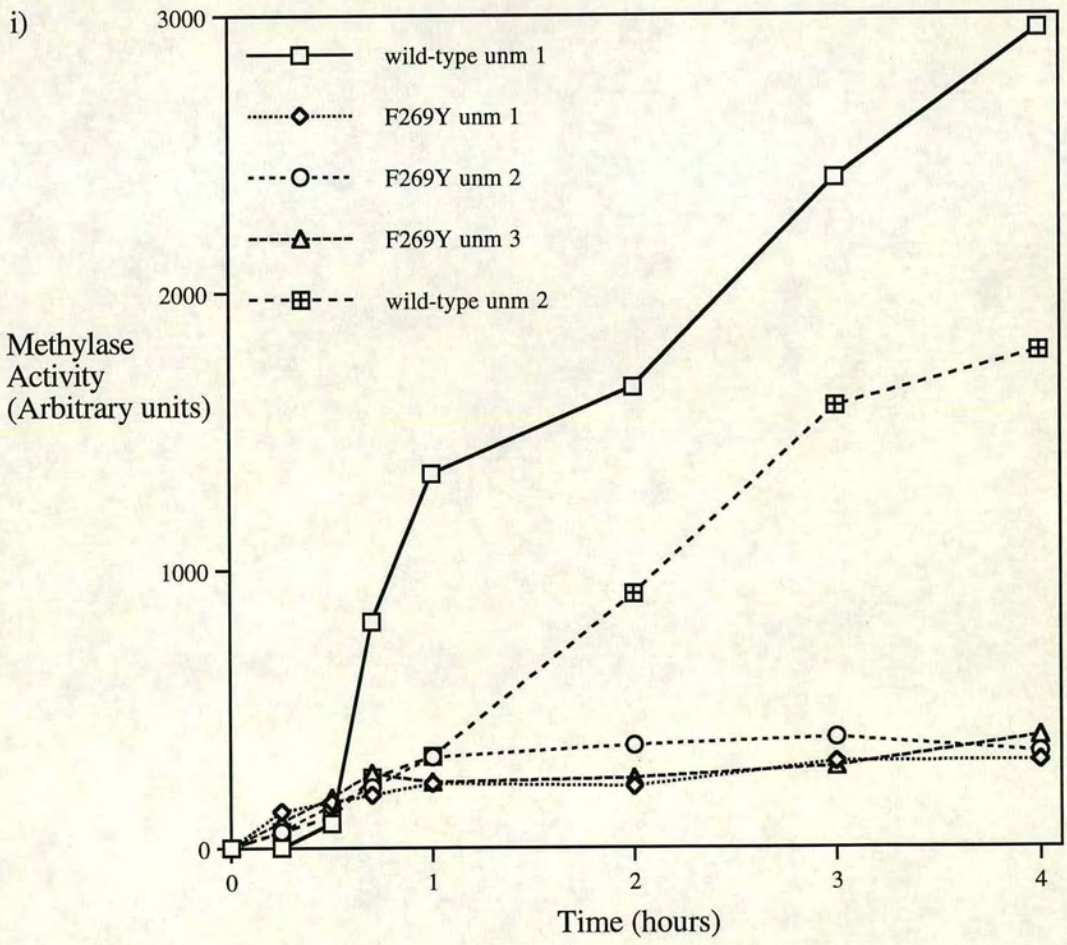


Figure 3.6 (2 of 3)

In vitro **Methylase Activity (fluorographic Detection of [*methyl* - ^{3}H] Transfer).**

ii) F269Y on hemimethylated DNA Substrate 860/409.

Figure 3.6 (2 of 3)

ii)

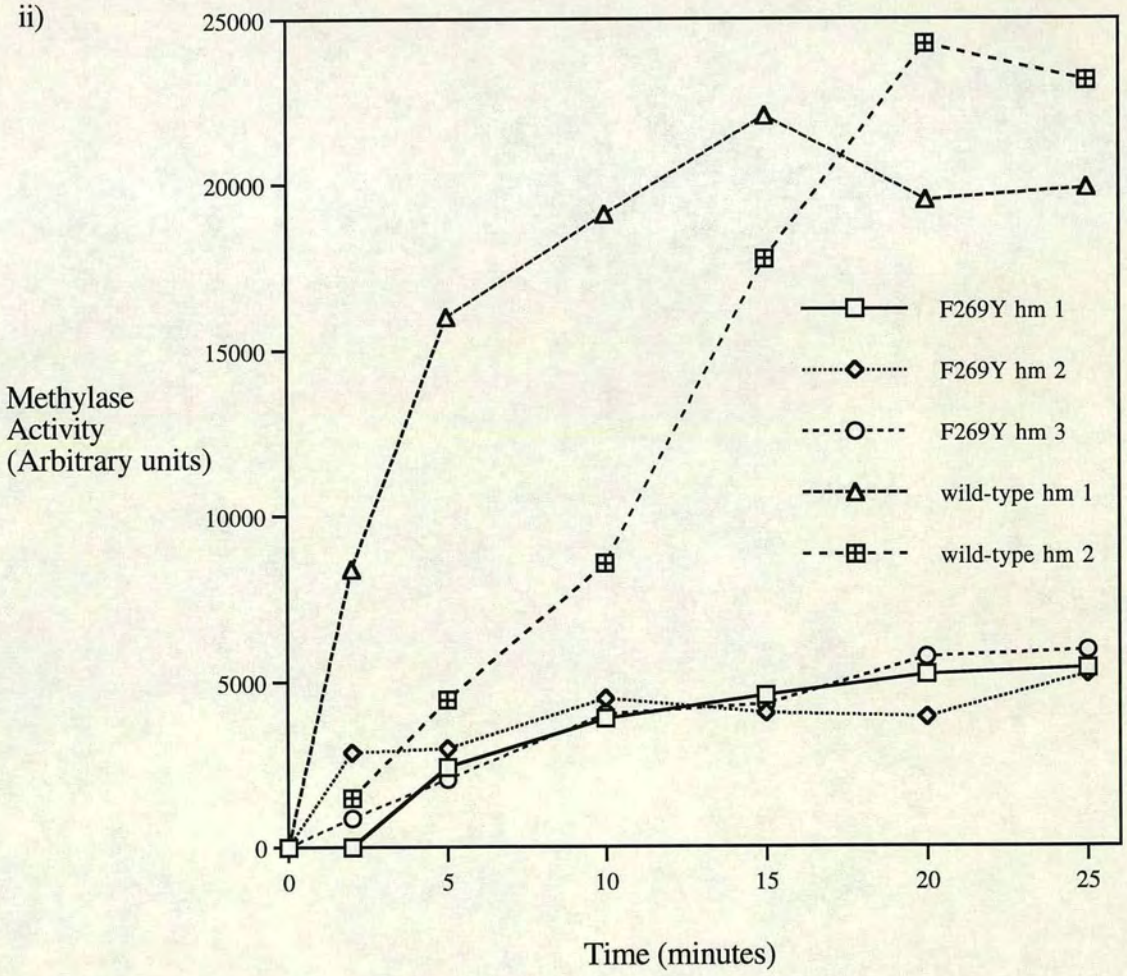


Figure 3.6 (3 of 3)

In vitro **Methylase Activity (fluorographic Detection of [*methyl* - ^{3}H] Transfer).**

iii) F269W on hemimethylated DNA Substrate 860/409.

Figure 3.6 (3 of 3)

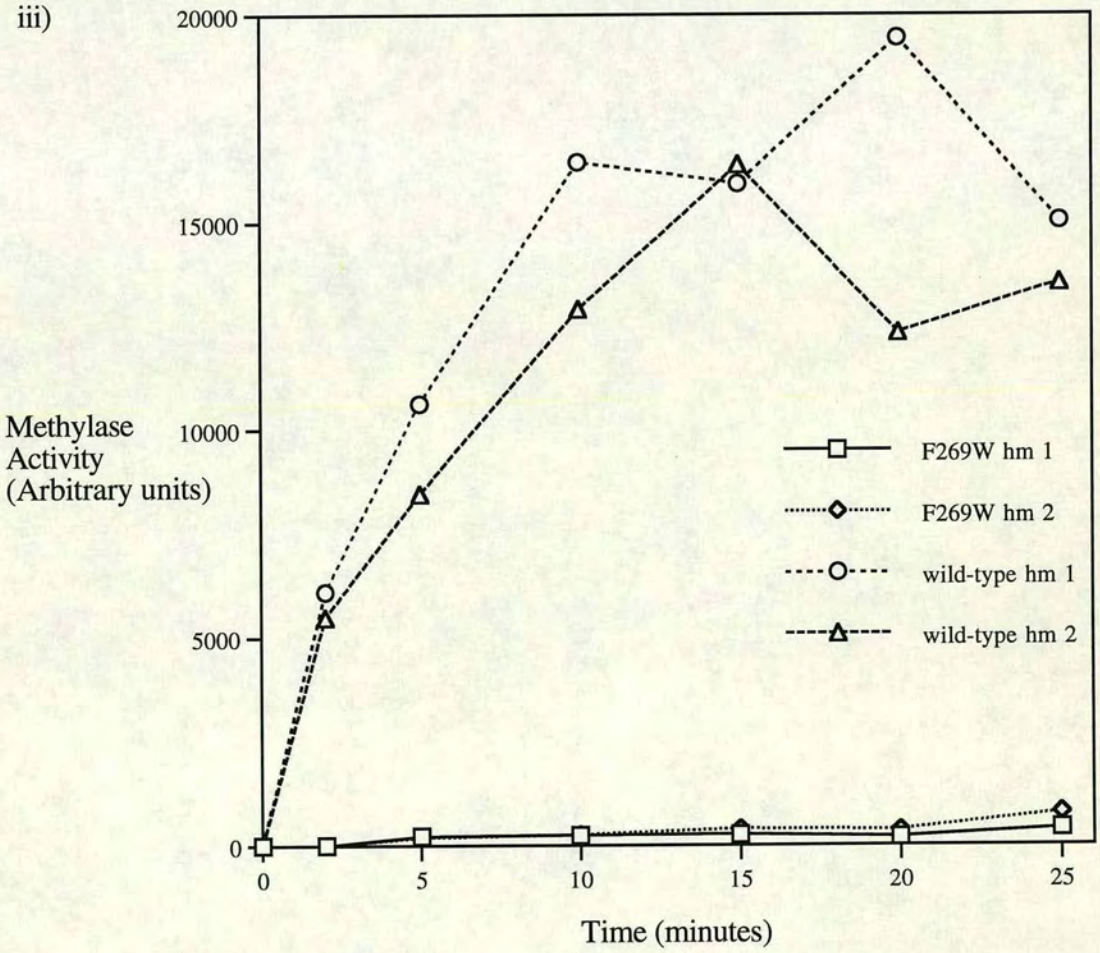


Table 3.5 Relative *in vitro* Methylase Activities of Mutants

Methylase	Relative Activity (% wild-type) on Hemimethylated Substrate (860/409)	Relative Activity (% wild-type) on Unmethylated Substrate (408/409)
G177D	0	0
F269C	0	0
F269G	0	0
N266D	0	0
F269Y	20-40	20-40
F269W	<5	<5*

*F269W may have some activity on an unmethylated substrate but separation of activity from background was not categorical at this low level of activity (figure 3.5 VI).

As in the measurement of *in vivo* activity, the mutants fall into three classes. All but two of the mutants fail to demonstrate any detectable activity on either an unmethylated or hemimethylated substrate (figure 3.5). G177D was assayed at both 25°C and 37°C and failed to show activity at either temperature (the assay of G177D in figure 3.5 was performed at 25°C, the result at 37°C being identical). There remains the possibility that these enzymes are not completely inactive but that their activities are below the detection level of the assay employed. However the mutants F269Y and F269W did demonstrate activity. F269Y has approximately 20-40% of wild-type activity on both the hemimethylated and unmethylated substrates. F269W has less than 5% of wild-type activity on a hemimethylated substrate and may possess a similar percentage activity on an unmethylated substrate although at this level, separation of activity from background by scanning of fluorographs was very difficult. Note that the exposure lengths for the F269W activity measurements are 23 days compared to the control wild-type exposure lengths of only 3 days (Figure 3.5). Exposure lengths for all other mutant enzyme assays are the same length as their respective wild-type control assays. Prolonged exposure of F269C, F269G, N269D and G177D assay gels have failed to indicate any activity (Dryden, D.T.F personal communication).

The *in vitro* methylase activity results support the *in vivo* activity results in that only F269Y and F269W have catalytic activity. At the *in vivo* level F269Y and the wild-type are indistinguishable but at the *in vitro* level F269Y displays a relatively decreased level of methylase activity. Presumably F269Y also has a decreased activity *in vivo* but this level is sufficient to provide the same protection to the host DNA as is afforded by the wild-type enzyme.

Of the residues tested at position 269, only tyrosine and tryptophan provide any semblance of *in vivo* or *in vitro* activity. This result indicates that the presence of an aromatic residue is required at this site in order to achieve enzyme activity.

3.6 Protein Denaturation by Guanidine Chloride

Denaturation by guanidine chloride (GuCl) was used in order to determine whether the stability of the mutants was significantly perturbed from that of the wild-type. One of the techniques for following the denaturation is to monitor fluorescence which in turn gives information about the environment of aromatic residues, particularly tryptophan (Lakowicz, 1983). A quantitative analysis can be used to determine by what difference in free energy the native regular conformation of a protein is stabilised compared to its unfolded state. A mutant of altered conformation is likely to demonstrate changes to this difference in free energy (Pace *et al.*, 1989).

The wild-type enzyme shows a distinct change in its intrinsic fluorescence when titrated with GuCl. The ratio of emission at 370nm to that at 320nm is approximately 0.5 at 0M GuCl, increasing to 3 at 4M GuCl. This red shift is characteristic of an increase in the solvent (H_2O) exposure of tryptophan residues, as one would expect in the denaturation of a protein. Furthermore the wild-type shows a characteristic peak in its anisotropy at 345nm between 0.7 and 1.2M GuCl. Anisotropy is a measure of the polarisation of the fluorescence emission. Any factor which affects the size, shape or the flexibility of a macromolecule will affect the observed anisotropy. The fluorescence at 320nm, 370nm, the fluorescence ratio (F370/F320) and the anisotropy at 345nm were examined for the following mutant proteins; G177D, F269C, F269G and F269Y (figure 3.7). G177D was chosen because its insolubility when purified at 37°C had indicated that it may have a reduced stability compared to the wild-type. F269C and F269G were chosen as examples of inactive mutants from region II and F269Y was chosen as an example of an active mutant from region II. The resulting graphs were compared to the wild-type controls in order to indicate any changes in stability (a quantitative thermodynamic analysis requires that the denaturation process is shown to be reversible and this was not attempted with the mutant proteins).

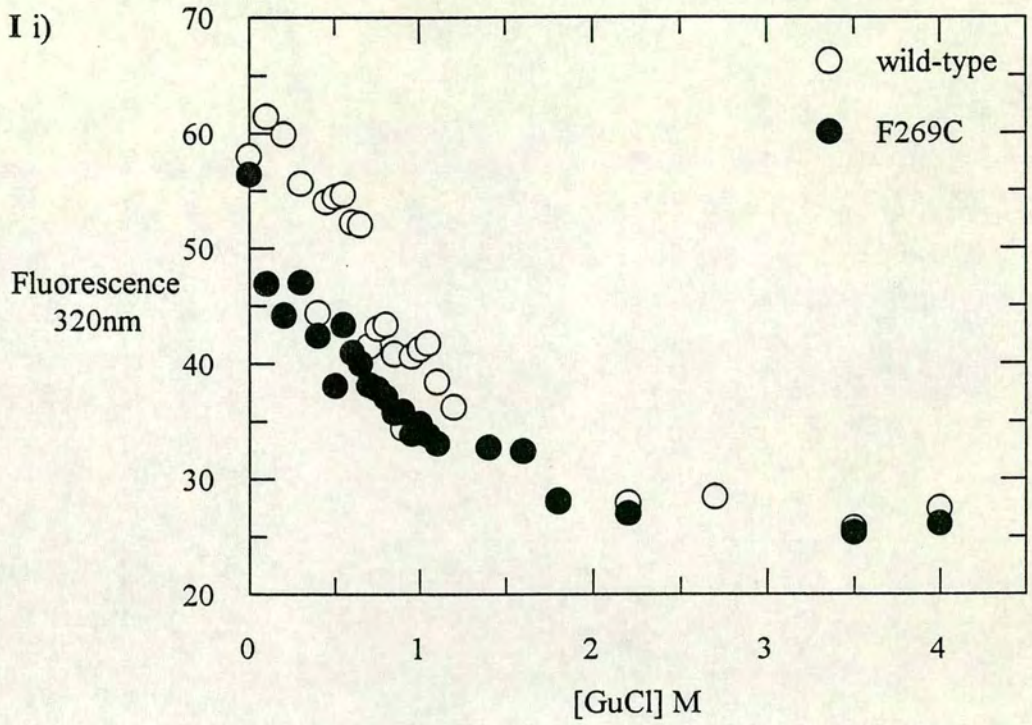
Figure 3.7 (1 of 8)

Protein Denaturation by Guanidine Chloride: F269C

I i) Fluorescence at 320nm vs. [GuCl]

I ii) Fluorescence at 370nm vs. [GuCl]

Figure 3.7 (1 of 8): F269C



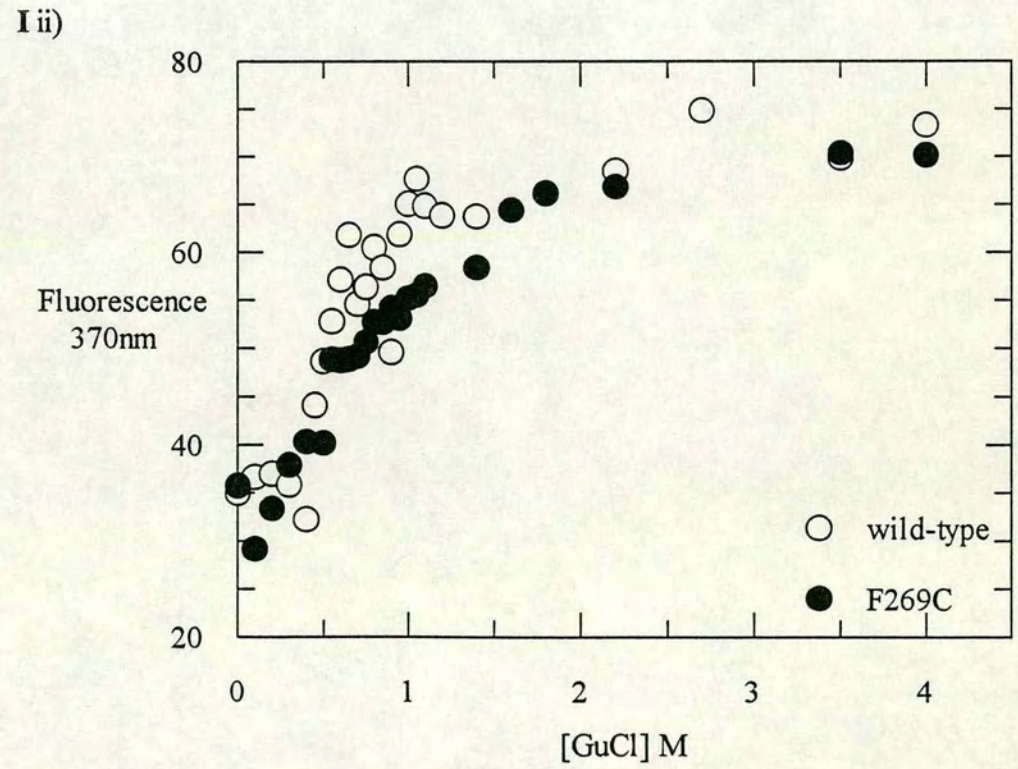


Figure 3.7 (2 of 8)

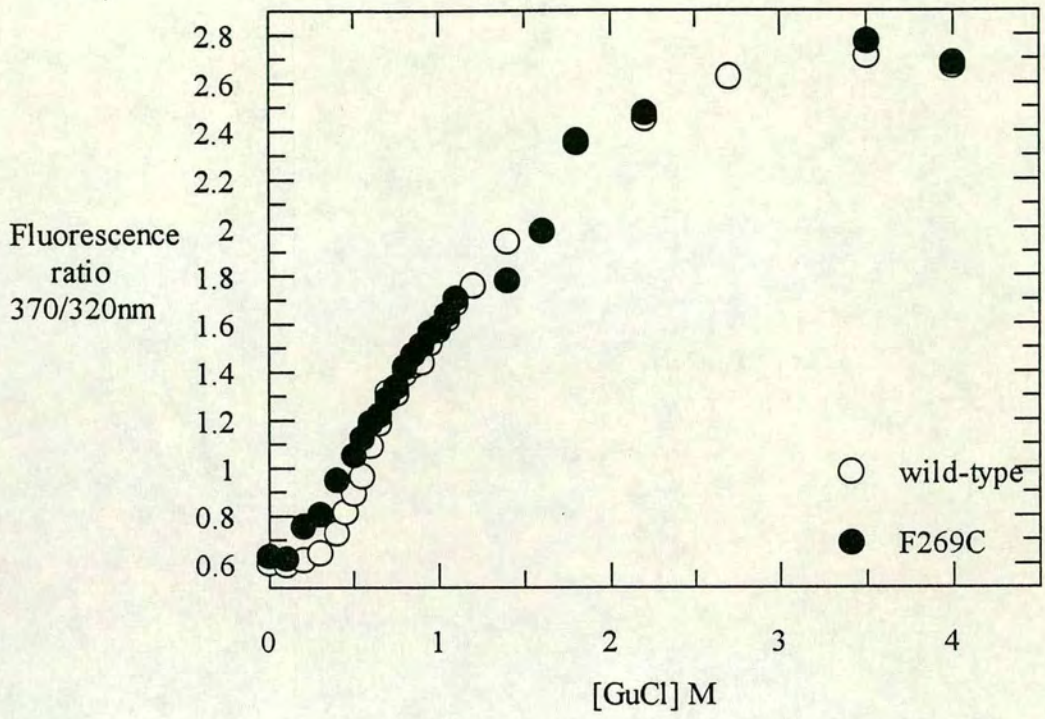
Protein Denaturation by Guanidine Chloride: F269C

I iii) Fluorescence ratio (370/320nm) vs. [GuCl]

I iv) Anisotropy (345nm) vs. [GuCl]

Figure 3.7 (2 of 8): F269C

I iii)



I iv)

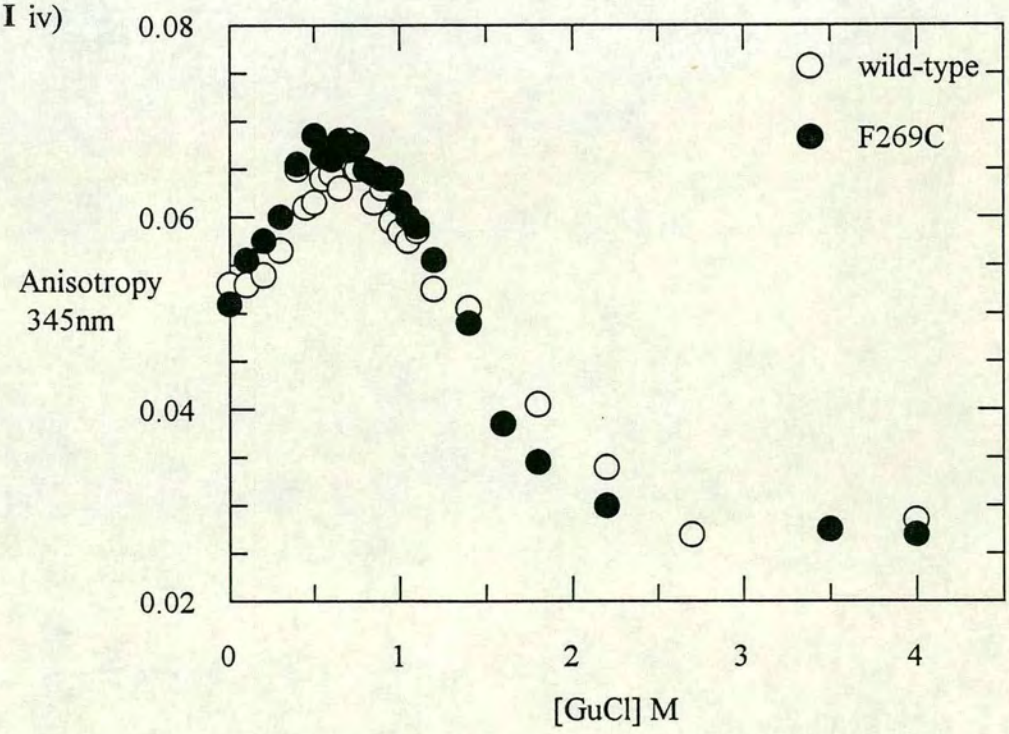


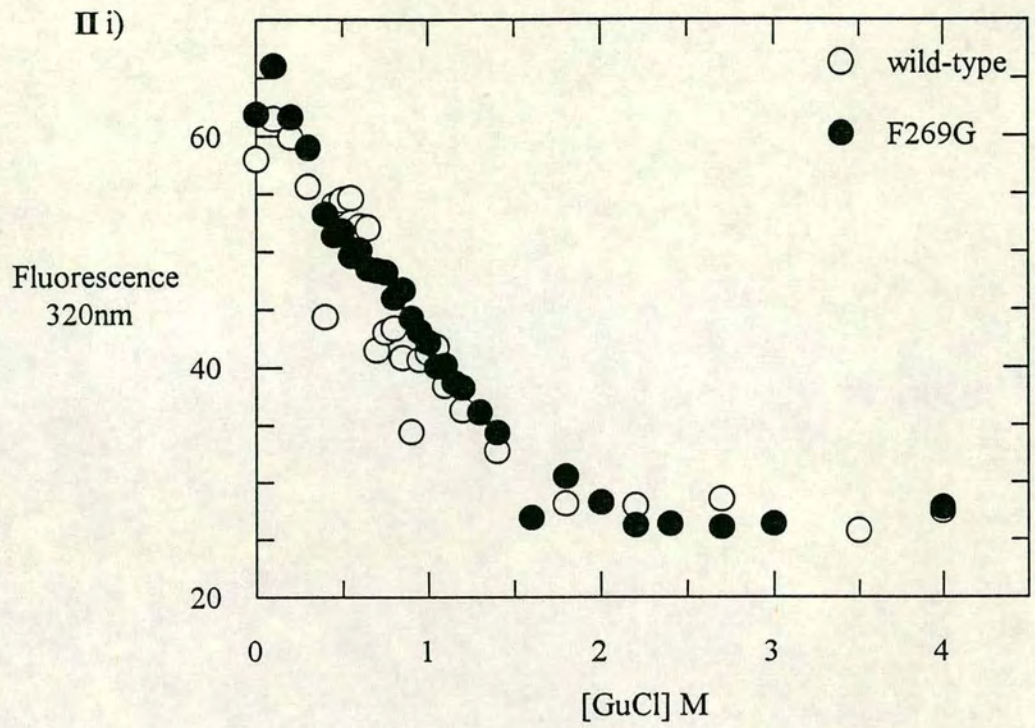
Figure 3.7 (3 of 8)

Protein Denaturation by Guanidine Chloride: F269G

II i) Fluorescence at 320nm vs. [GuCl]

II ii) Fluorescence at 370nm vs. [GuCl]

Figure 3.7 (3 of 8): F269G



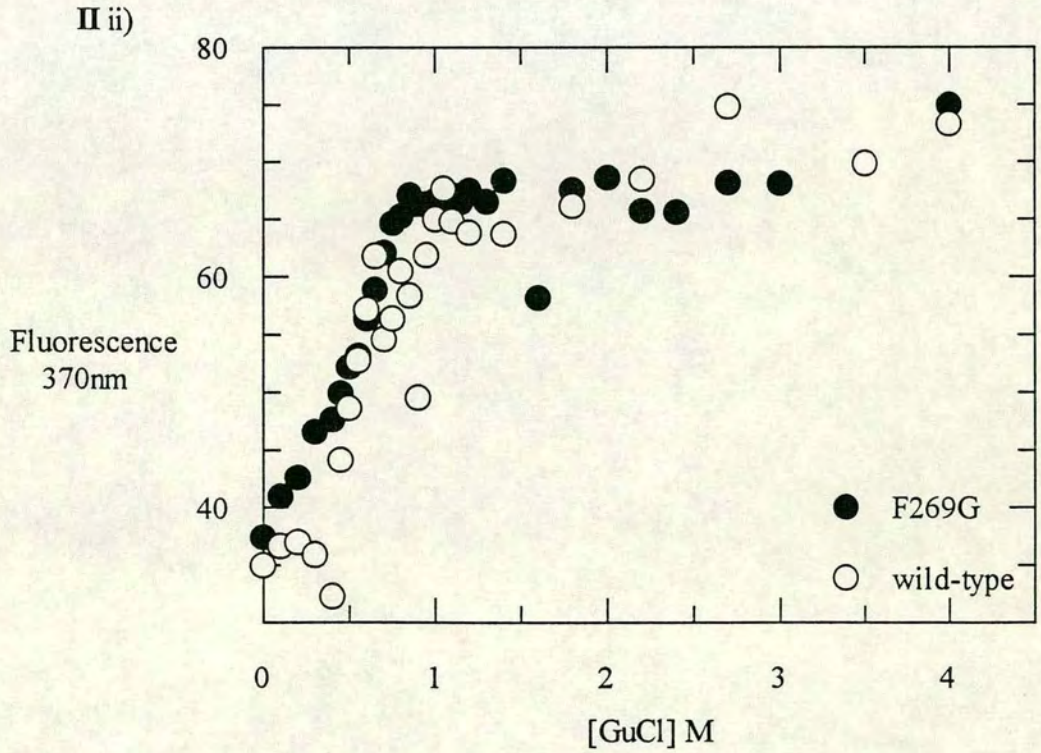


Figure 3.7 (4 of 8)

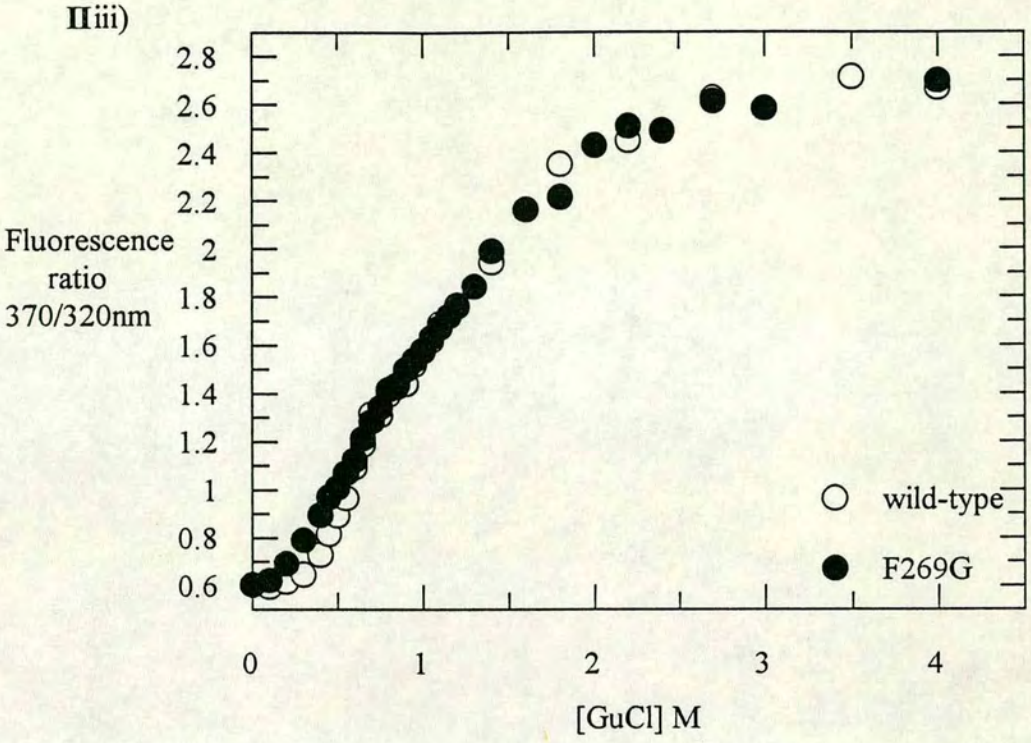
Protein Denaturation by Guanidine Chloride: F269G

II iii) Fluorescence ratio (370/320nm) vs. [GuCl]

II iv) Anisotropy (345nm) vs. [GuCl]

Figure 3.7 (4 of 8): F269G

IIiii)



IIiv)

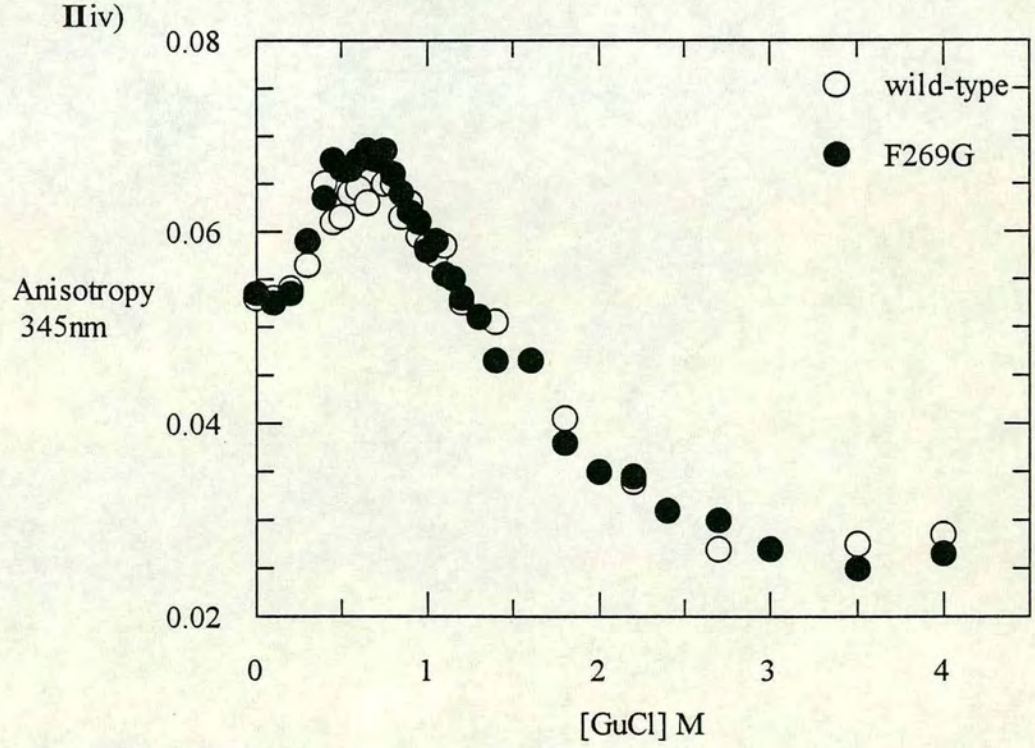


Figure 3.7 (5 of 8)

Protein Denaturation by Guanidine Chloride: F269Y

III i) Fluorescence at 320nm vs. [GuCl]

III ii) Fluorescence at 370nm vs. [GuCl]

Figure 3.7 (5 of 8): F269Y

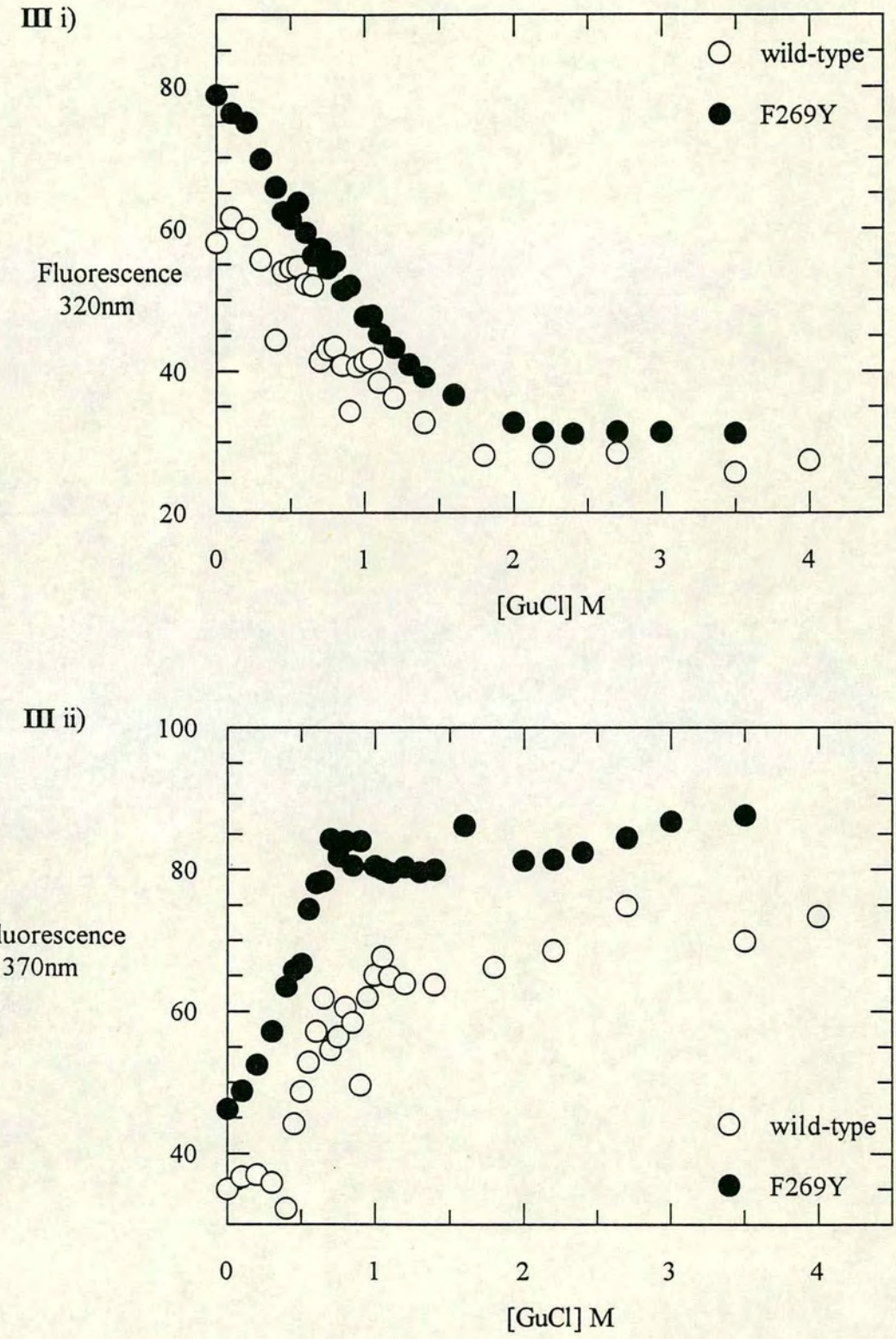


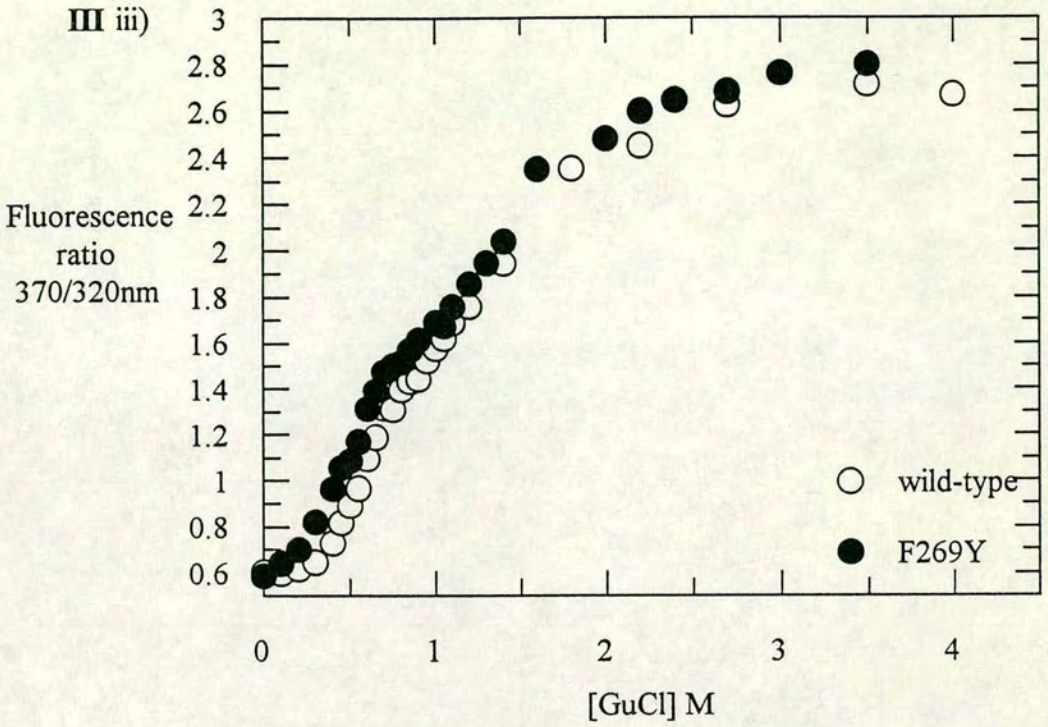
Figure 3.7 (6 of 8)

Protein Denaturation by Guanidine Chloride: F269Y

III iii) Fluorescence ratio (370/320nm) vs. [GuCl]

III iv) Anisotropy (345nm) vs. [GuCl]

Figure 3.7 (6 of 8): F269Y



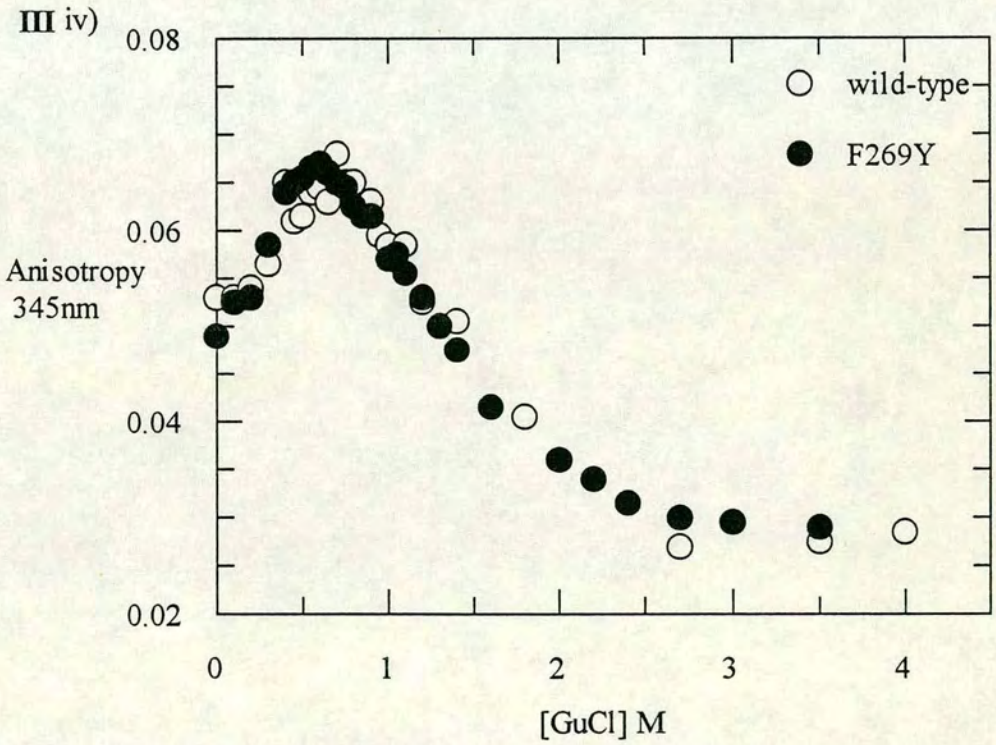


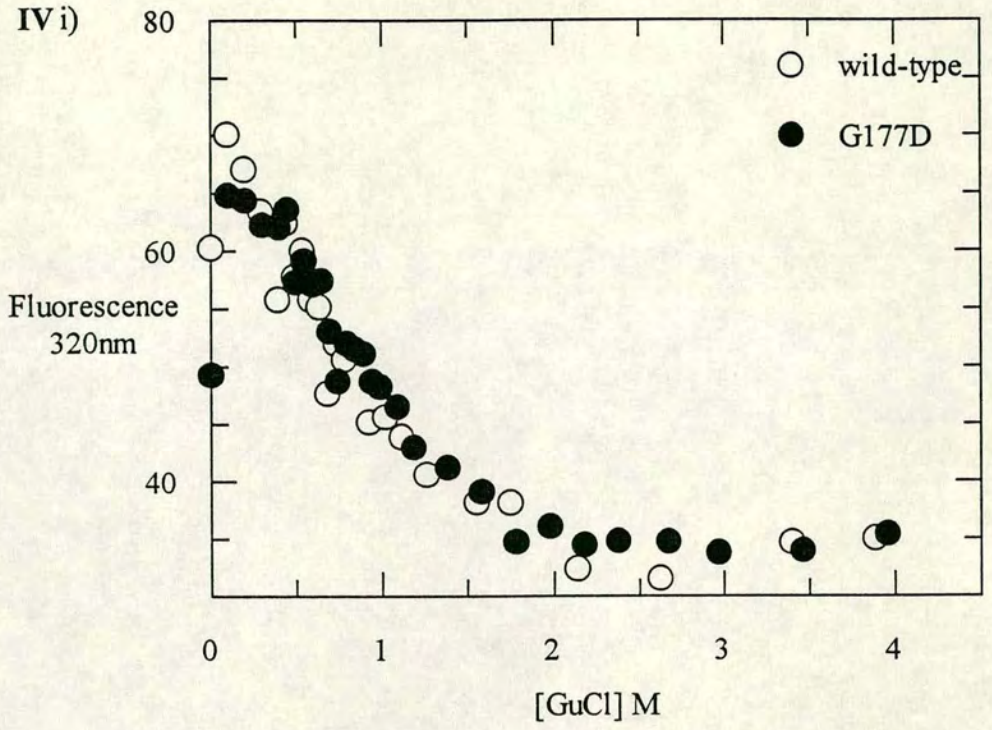
Figure 3.7 (7 of 8)

Protein Denaturation by Guanidine Chloride: G177D

IV i) Fluorescence at 320nm vs. [GuCl]

IV ii) Fluorescence at 370nm vs. [GuCl]

Figure 3.7 (7 of 8): G177D



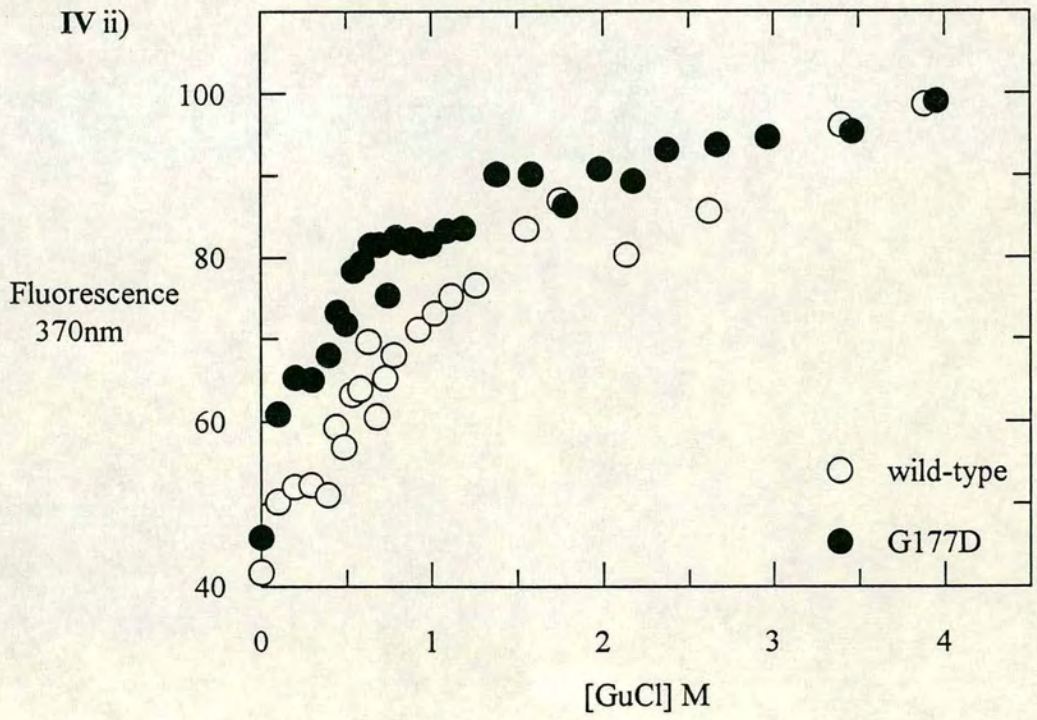


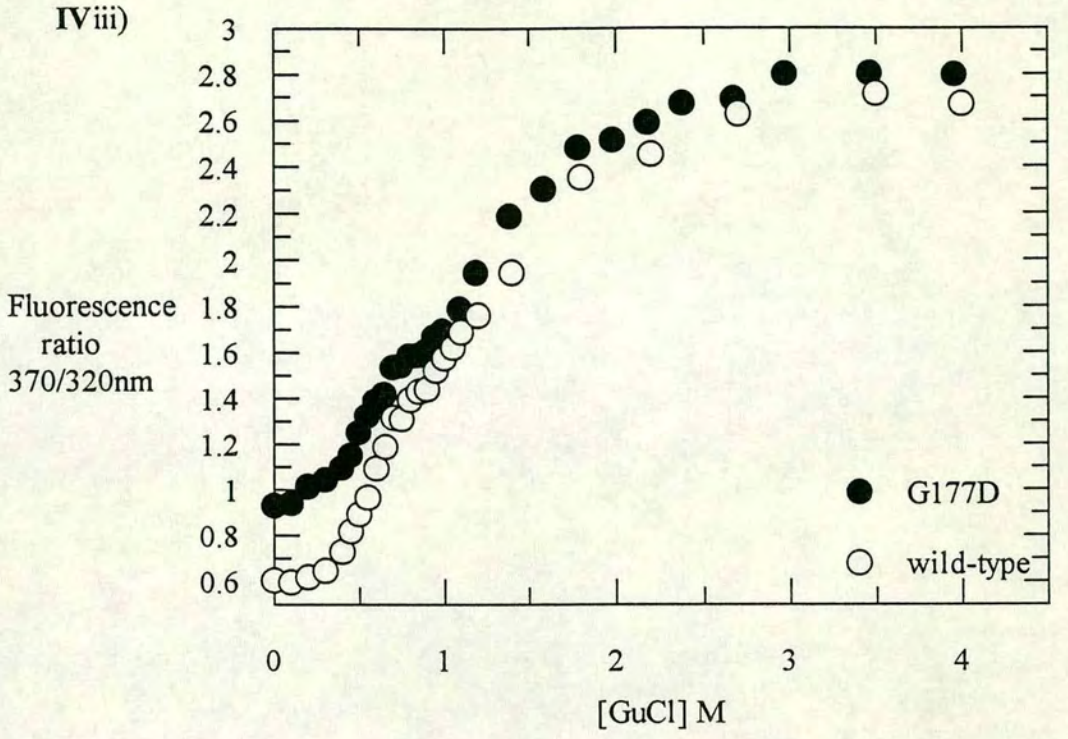
Figure 3.7 (8 of 8)

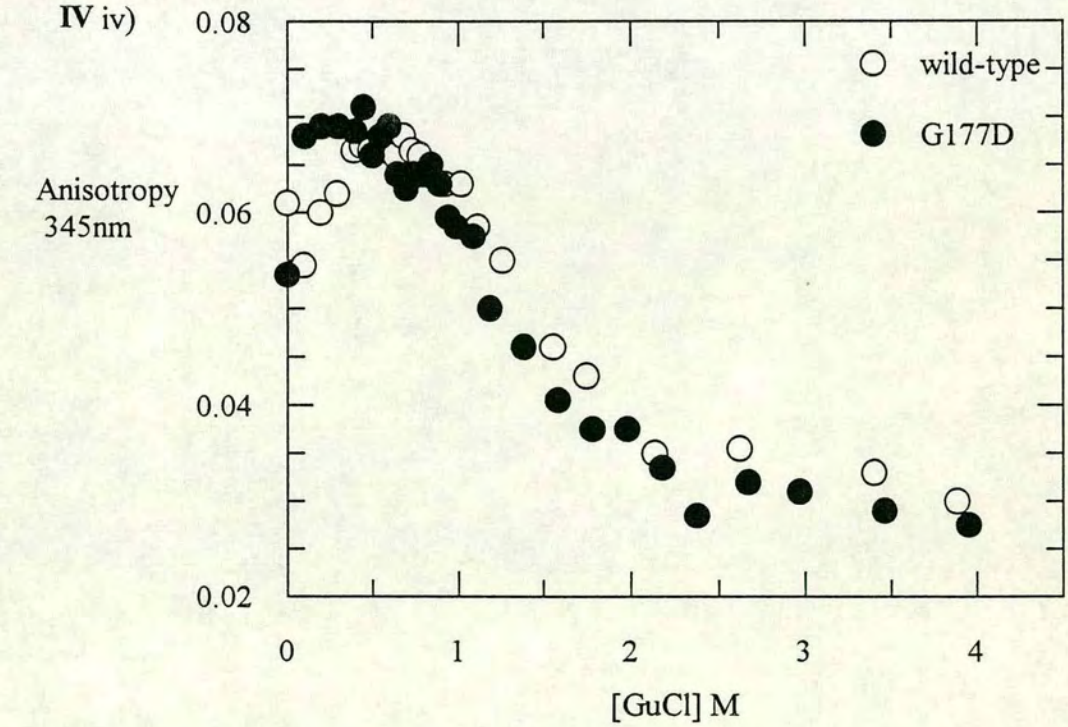
Protein Denaturation by Guanidine Chloride: G177D

IV iii) Fluorescence ratio (370/320nm) vs. [GuCl]

IV iv) Anisotropy (345nm) vs. [GuCl]

Figure 3.7 (8 of 8): G177D





In the same manner as the wild-type, all the mutant proteins inspected display a characteristic increase in anisotropy at approximately 0.5M GuCl and a subsequent decrease at approximately 1M GuCl. The increase in anisotropy has been attributed, in the wild-type, to one or more of the following effects (Dryden *et al.*, 1993);

1) The formation of aggregates resulting in an increase in the size of the fluorescent body which in turn restricts fluorophore motion.

2) A reduction in fluorescence energy transfer (and therefore an increase in polarisation) as a result of subunit separation.

3) Light scattering from aggregates.

Furthermore, the mutant proteins tend to display a transition in fluorescence at approximately 0.5M GuCl, when measured at 370nm, followed by a plateau in intensity which is complete at 1.5M GuCl. However, the transition as followed by anisotropy or the fluorescence ratio (F370/F320) is not complete until 3M GuCl. The same effects are found in the wild-type protein. Dryden *et al.*, 1993, have interpreted the first transition (complete at 1.5M GuCl, as followed by fluorescence at 370nm) as the transition between the native protein and an intermediate state. The second transition, complete at 3.5M GuCl (as followed by anisotropy or the fluorescence ratio, F370/F320) was interpreted as the transition between an intermediate state and the unfolded state.

The observed anisotropy, fluorescence and the fluorescence ratio suggest that the four mutants analysed do not have any significant deviation from the wild-type in their stability. In the case of G177D (figure 3.7 IV iv) the increase in anisotropy may occur at slightly lower GuCl concentration than for the wild-type but this is unlikely to have a significant effect upon stability. This implies that the inactivity of the F269C, F269G and G177D enzymes is probably not due to either a major conformational change in HsdM or failure of the HsdM and S subunits to interact correctly but is more likely the result of some local alteration. Furthermore, the instability of G177D when expressed at 37°C is not evident at 20°C, the temperature at which the denaturation experiments were performed. Hence G177D appears to have a similar stability as the wild-type at 20°C and its inactivity at this temperature appears more likely to be due to a local alteration in protein structure.

The denaturation curves of G177D suggest that any conformational change caused by the mutation is not large scale and if present is a local alteration to the region I motif. The insolubility of this mutant at 37°C (section 3.4) perhaps suggests

that at this higher temperature its folding pathway is altered such that, for example, an intermediate in the pathway is more unstable and results in insolubility. At 25°C this effect is not as apparent and the folding pathway can proceed as far as the final, fully folded conformation.

3.7 DNA Binding

DNA binding by the mutant proteins was investigated using gel retardation. This technique has been used to estimate the dissociation constants (K_d) of several DNA substrates for the wild-type enzyme. The wild-type methylase shows tighter binding for a specific DNA substrate (i.e. one that contains a target recognition site) than for a non-specific DNA substrate. In addition, binding to both substrates is increased in the presence of the cofactor AdoMet. Furthermore, the K_d values for unmethylated and hemimethylated substrates are similar indicating that the strong specificity of the wild-type for a hemimethylated DNA does not have its basis in a difference in binding affinity (Powell *et al.*, 1993).

The preferred substrate for methylation is hemimethylated DNA and it is likely that when using this substrate in a DNA binding assay, the majority of sites will be methylated to give fully methylated DNA. In order to investigate whether the presence of the resulting fully methylated DNA can affect the K_d value, Powell *et al.* (1993) performed the experiments at both 22°C and 4°C (at 4°C, methylation would be expected to proceed much more slowly). It was found that the K_d values observed were the same at both temperatures and hence the methylation of a hemimethylated substrate had no discernible effect on binding.

In this study, binding of a hemimethylated (409/860) and a non-specific (862/863) DNA substrate was investigated for each of the mutants (table 3.7). The hemimethylated duplex was chosen as an example of a specific substrate and its binding by the mutants was compared to the binding of the non-specific substrate. In addition, the effect of AdoMet on the binding of both substrates was examined.

Table 3.7 Oligonucleotide Duplexes used in DNA binding

Oligonucleotide	Oligonucleotide Sequence
Specific Substrate	
860/ 409	5'-TGTCTAGATATCGGCCTA^{CH_3}**AAC**CACGTG**GTGC**GTACGAGCTCAGGCG ACAGATCTATAGCCGGA**TTG**GTGCAC**CACG**CATGCTCGAGTCCGC-5'
Non-specific Substrate	
862/ 863	5'-TGTCTAGATATCGGCCT**CCA**CACGTG**TGTA**GTACGAGCTCAGGCG ACAGATCTATAGCCGGA**GGT**GTGCAC**ACAT**CATGCTCGAGTCCGC-5'

As with the wild-type enzyme, increasing titration of DNA substrate with protein resulted in a decrease in free DNA and an increase in higher molecular weight DNA-protein complexes (figure 3.8). The first complex to appear has been assumed to be of the ratio 1:1 (Powell *et al.*, 1993) and for each of the mutant proteins is more prominent when using a specific DNA substrate. At increased protein concentrations a slower migrating complex appears and, as with the wild-type enzyme, this species is more prominent when using a non-specific substrate. This second complex is not as well resolved as the 1:1 complex perhaps because of an increased propensity to dissociation within the gel. Alternatively this second complex may be a combination of a series of closely migrating complexes which differ in the site of addition of subsequent protein molecules to the 1:1 complex.

Figure 3.8 (1 of 7)

DNA binding by Methylase as determined by Gel Retardation. The wild-type, F269C, F269G, F269W, F269Y, and N266D methylases show stronger binding to the specific substrate (860/409) than to the non-specific substrate (862/863). Furthermore, DNA binding by these methylases is enhanced in the presence of AdoMet. However, although G177D methylase appears to discriminate between the specific and non-specific substrates, the binding of neither substrate is enhanced by AdoMet.

I i) wild-type binding of specific DNA substrate 860/409 + AdoMet
ii) " " " " " " " - AdoMet
iii) wild-type binding of non-specific DNA substrate 862/863 + AdoMet
iv) " " " " " " " - AdoMet

Figure 3.8 DNA binding (1 of 7): wild-type

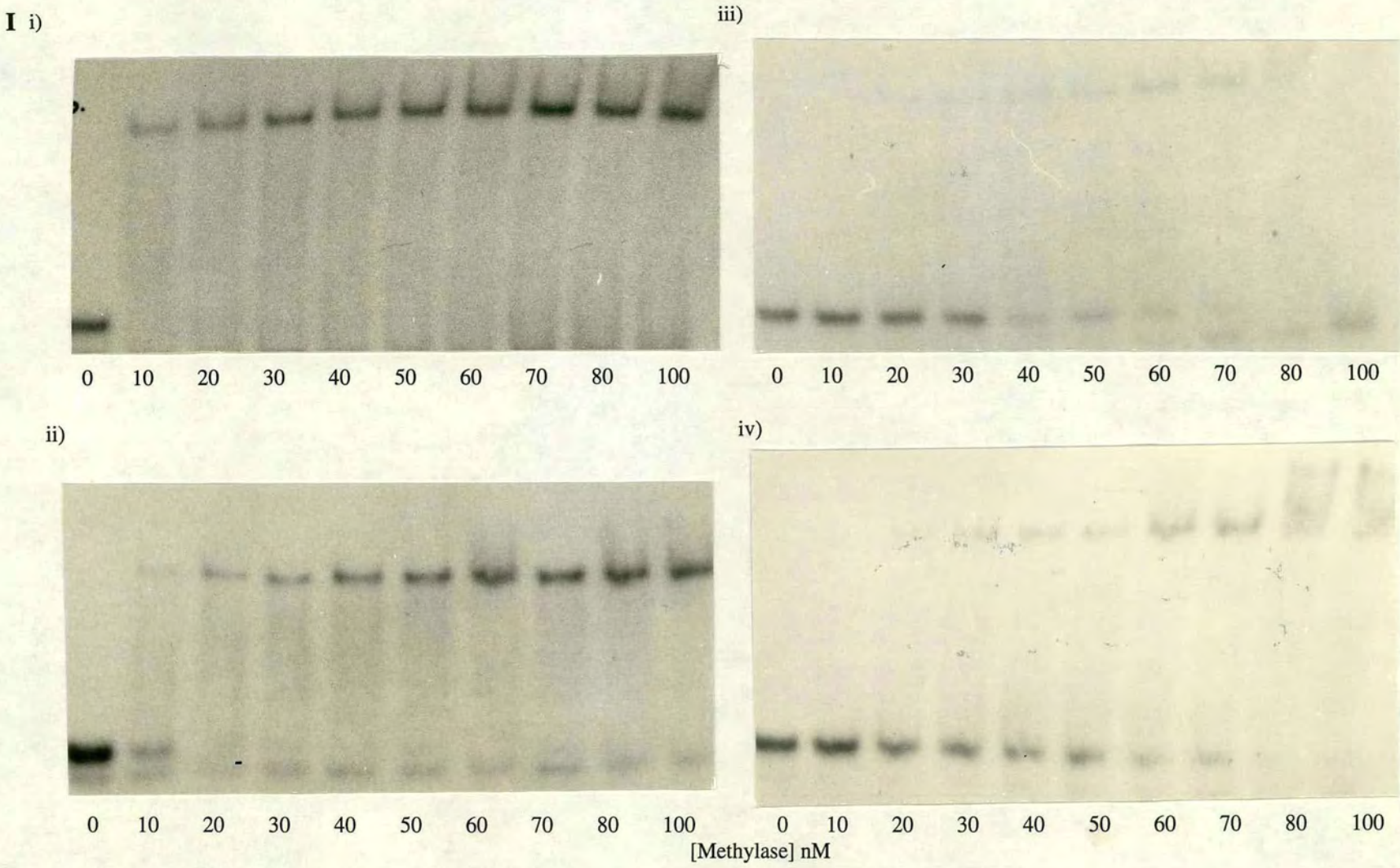


Figure 3.8 (2 of 7)

DNA binding by Methylase as determined by Gel Retardation.

II i) F269C binding of specific DNA substrate 860/409 + AdoMet

ii) " " " " " " " - AdoMet

iii) F269C binding of non-specific DNA substrate 862/863 + AdoMet

iv) " " " " " " " - AdoMet

Figure 3.8 DNA binding (2 of 7): F269C

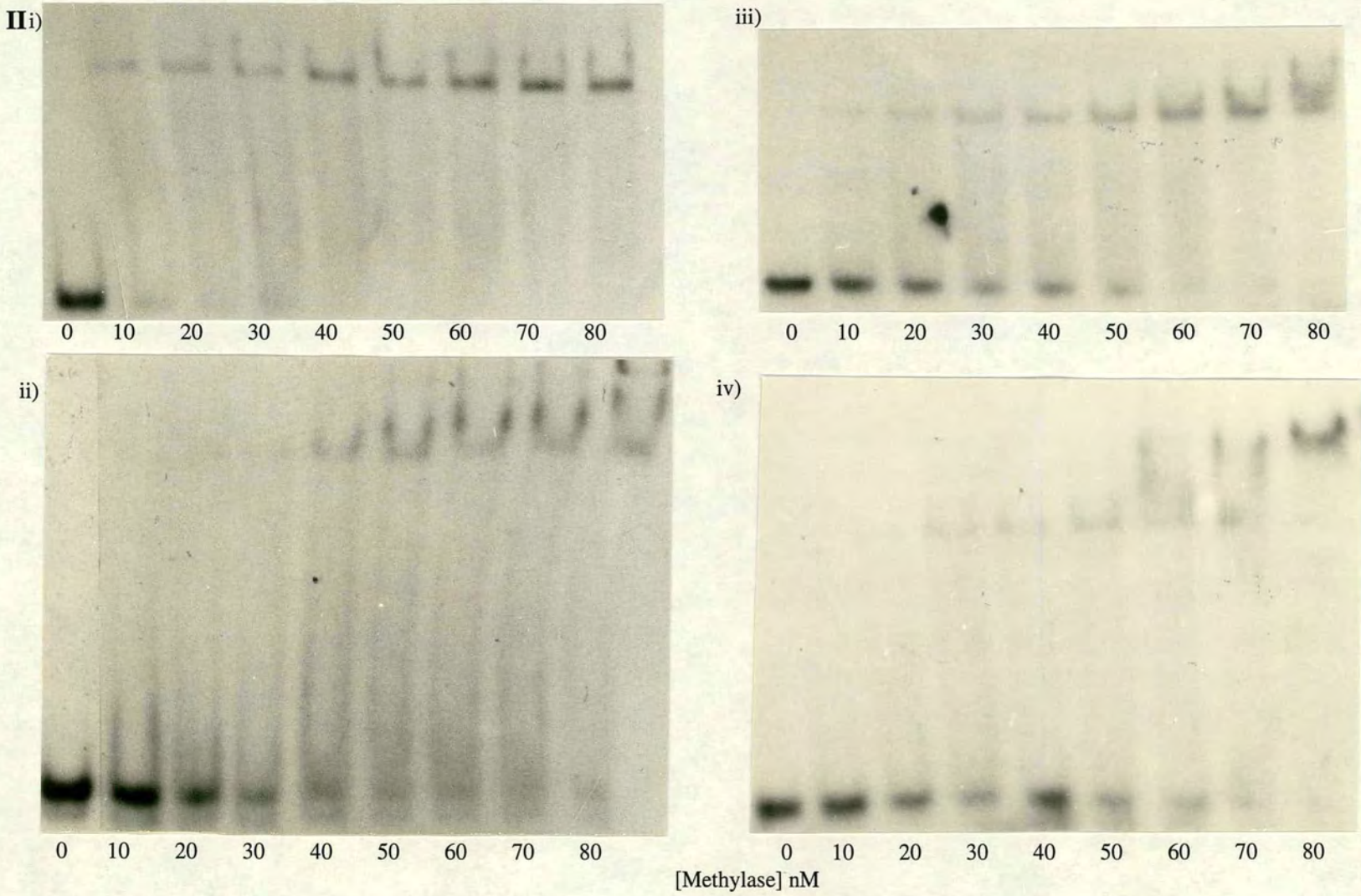


Figure 3.8 (3 of 7)

DNA binding by Methylase as determined by Gel Retardation.

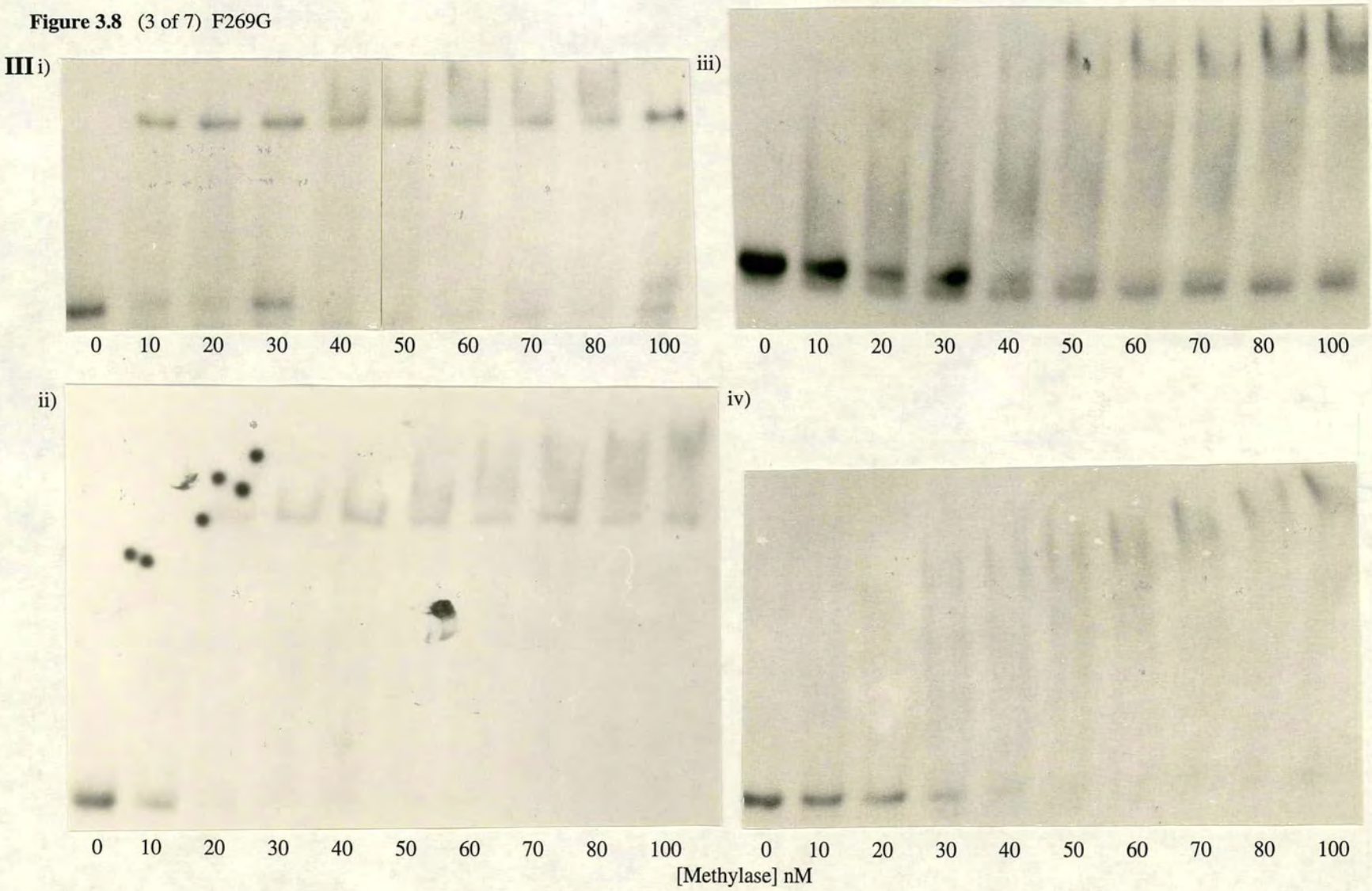
III i) F269G binding of specific DNA substrate 860/409 + AdoMet

ii) " " " " " - AdoMet

iii) F269G binding of non-specific DNA substrate 862/863 + AdoMet

iv) " " " " " " - AdoMet

Figure 3.8 (3 of 7) F269G



101a

Figure 3.8 (4 of 7)

DNA binding by Methylase as determined by Gel Retardation.

IV i) F269W binding of specific DNA substrate 860/409 + AdoMet

ii) " " " " " " " - AdoMet

iii) F269W binding of non-specific DNA substrate 862/863 + AdoMet

iv) " " " " " " " - AdoMet

Figure 3.8 DNA binding (4 of 7): F269W

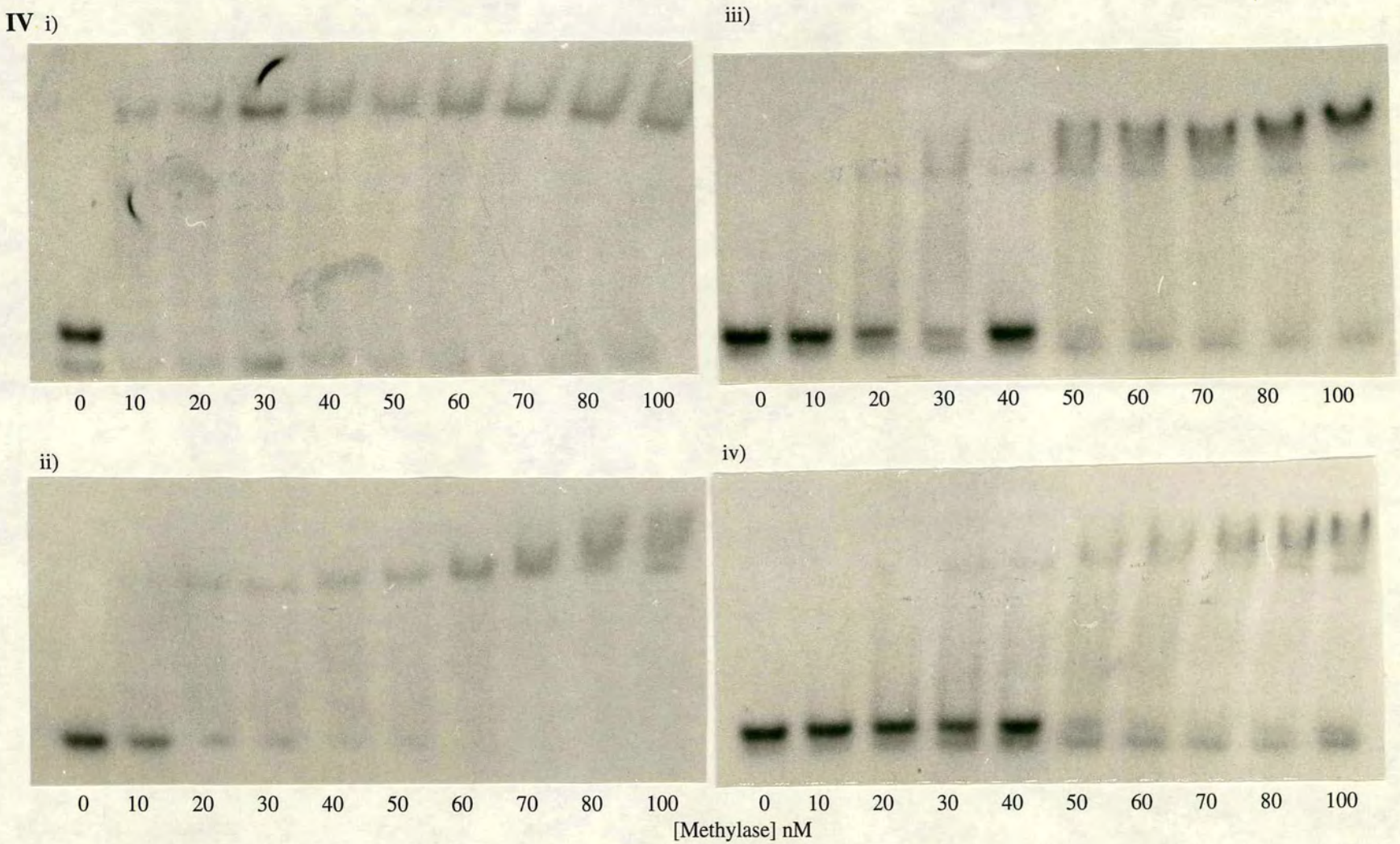


Figure 3.8 (5 of 7)

DNA binding by Methylase as determined by Gel Retardation.

V i) F269Y binding of specific DNA substrate 860/409 + AdoMet

ii) " " " " " " " - AdoMet

iii) F269Y binding of non-specific DNA substrate 862/863 + AdoMet

iv) " " " " " " " - AdoMet

Figure 3.8 DNA binding (5 of 7): F269Y

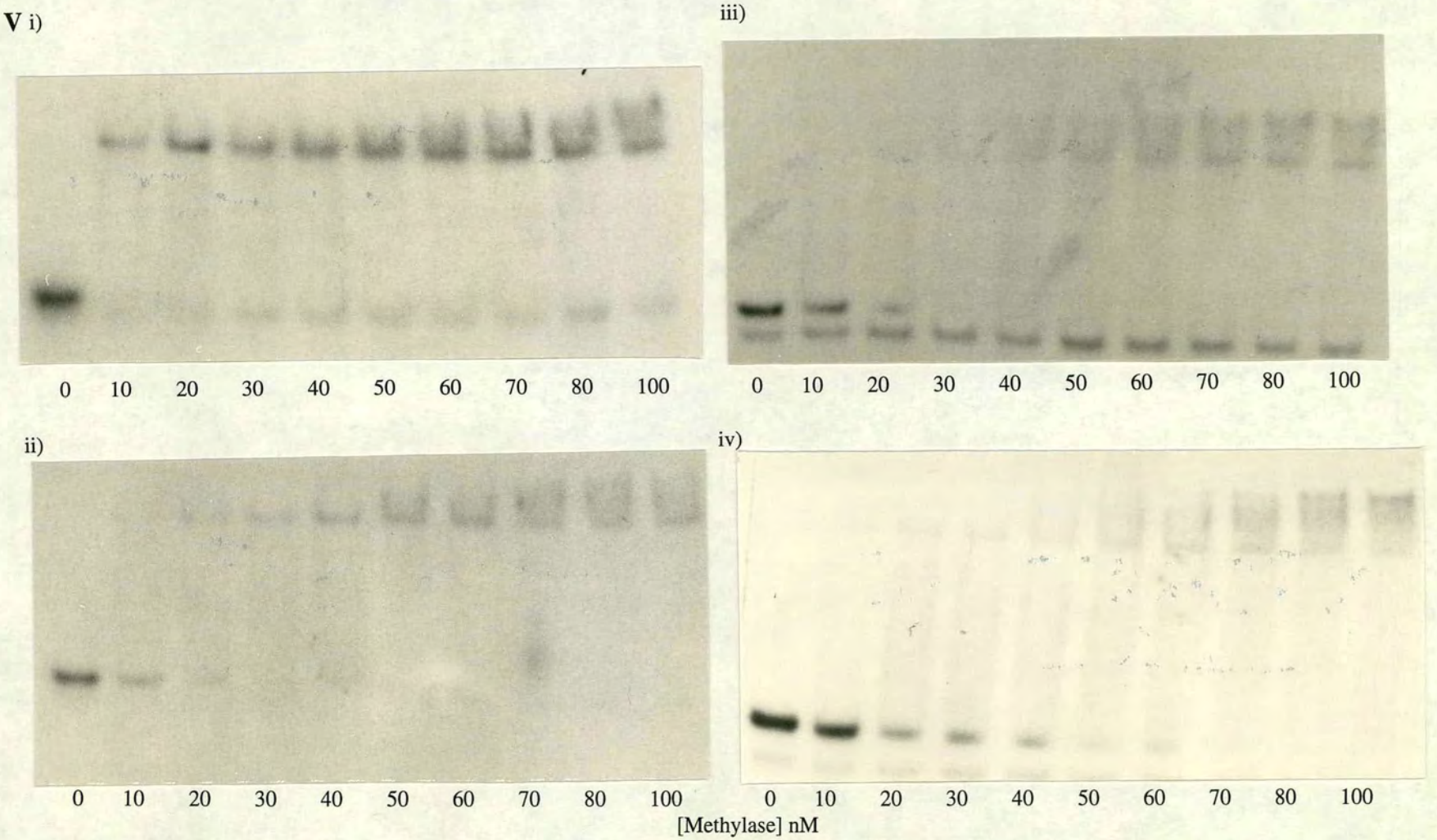


Figure 3.8 DNA binding (6 of 7): N266D

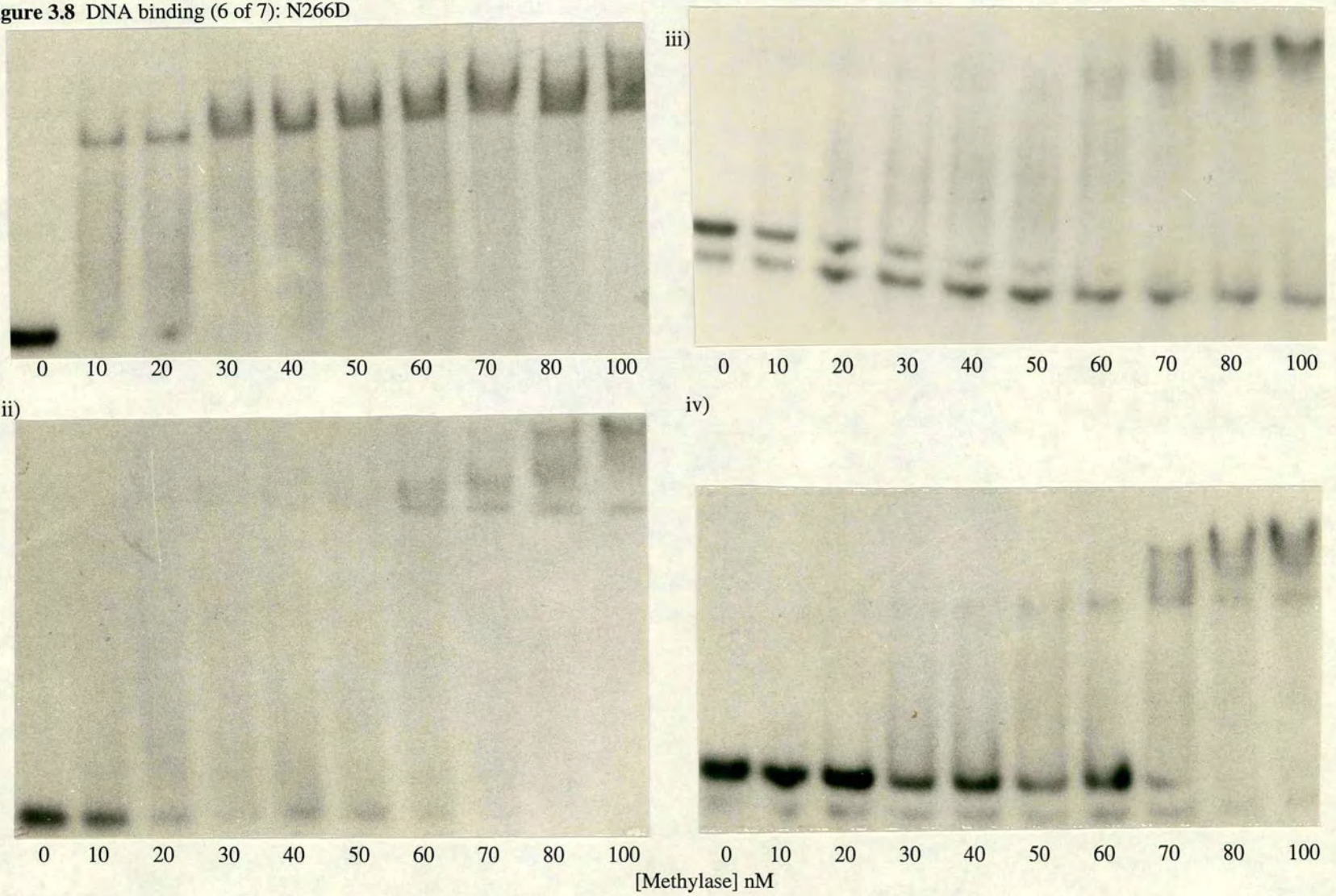


Figure 3.8 (7 of 7)

DNA binding by Methylase as determined by Gel Retardation.

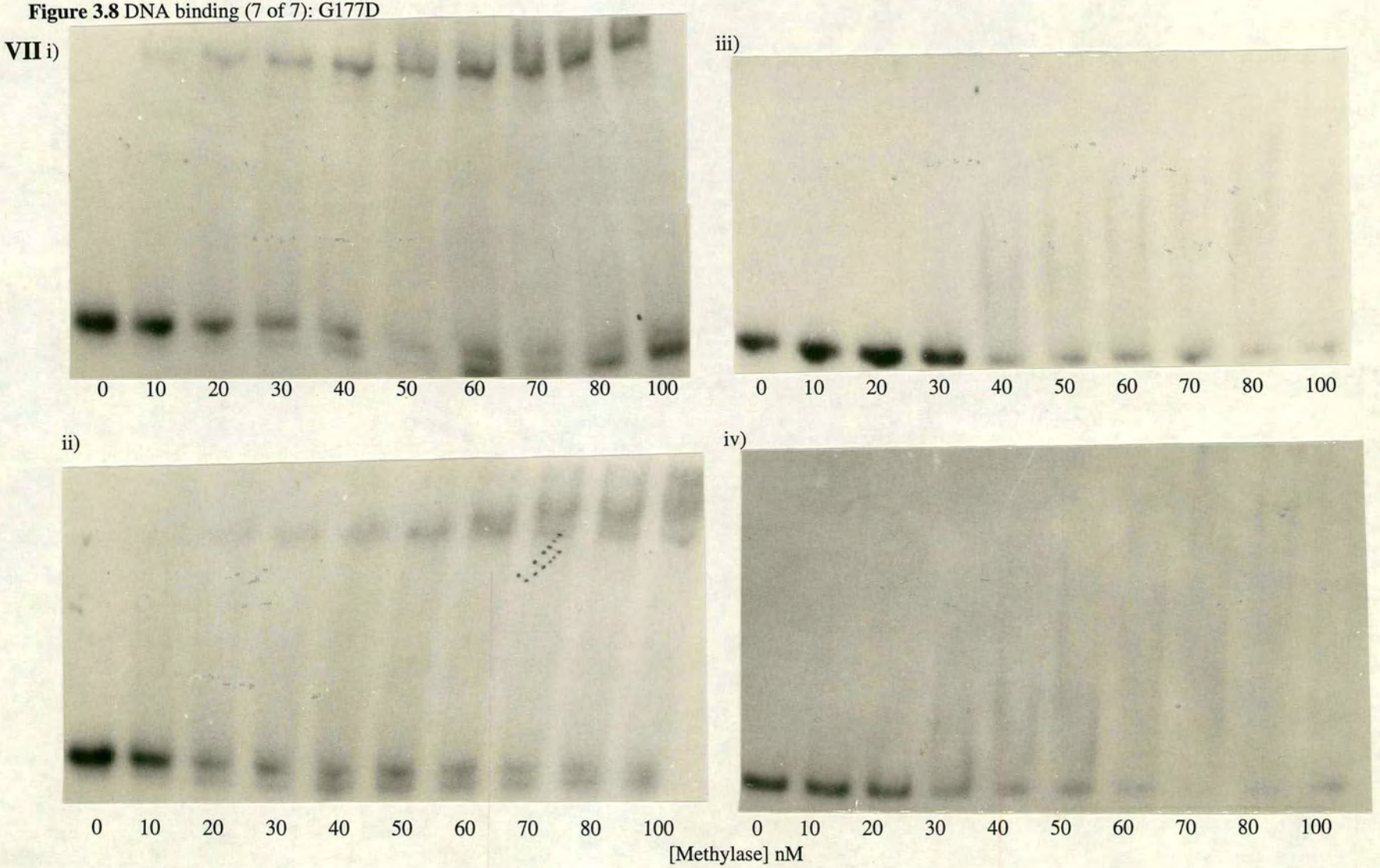
VII i) G177D binding of specific DNA substrate 860/409 + AdoMet

ii) " " " " " " " - AdoMet

iii) G177D binding of non-specific DNA substrate 862/863 + AdoMet

iv) " " " " " " " - AdoMet

Figure 3.8 DNA binding (7 of 7): G177D



Determination of Dissociation Constant (K_d) from Gel Retardation Data

In the simplest case a DNA binding protein (P) exists in only two states, bound to a specific DNA site (D) or free in solution. The following equilibrium is established:

$$D + P \overset{K_d}{\rightleftharpoons} DP$$

$$\text{Hence } K_d = \frac{[D][P]}{[DP]}$$

From the above equation, when [D]/[DP] = 1, K_d = [P]. Hence the dissociation constant, K_d, is the protein concentration at which free DNA is equal in concentration to bound DNA. If the protein concentration is in large excess over the DNA concentration, then the DNA-bound protein represents only a small percentage of the total protein and the approximation $[P] \sim [P_{total}]$ can be made (Fried, 1989). In the study of the wild-type enzyme it was found that if in binding experiments the DNA concentration is above 2.5nm, then an overestimate of the K_d is obtained. However, the K_d values obtained when using DNA concentrations below 2.5nm are all approximately equal (Powell *et al.*, 1993). For this reason a DNA concentration of 1nm was used in this study.

Measurement of free DNA provides the most accurate determination of complex formation. If a protein-DNA complex dissociates during migration within the gel, the released DNA will not catch up with the 'free DNA'. Therefore, as long as the lifetime of a complex exceeds the time required for free DNA to leave the sample well, the free DNA band represents the concentration of unbound DNA in the original solution. Hence the equation, $DNA_{bound} = DNA_{total} - DNA_{free}$, can be used even if complexes are not represented as distinct bands within the gel (Fried, 1989).

Table 3.7 Relative Dissociation Constants (K_d) for Specific and Non-specific DNA Substrates [a]

Methylase	Specific Substrate (409/860)				Non-specific Substrate (862/863)			
	-AdoMet	Mean	+AdoMet	Mean	-AdoMet	Mean	+AdoMet	Mean
wild-type[b]	9±2[b]		1.6±0.7[b]		90±15[b]		43±7[b]	
F269C	i)20±4[#] ii)16±5 iii)22±6	19±3	i)<5[#] ii)<5	<5	i)40±6[#] ii)36±4	38±5	i)30±8[#] ii)28±4	29±6
wt control	14±2		<5		40±8		30±5	
F269G	i)8±1[#] ii)10±3	9±2	i)<5[#] ii)<5	<5	i)24±2[#] ii)29±2	26.5±2	i)22±2[#] ii)18±1	20±1.5
wt control	12±1		<5		25±2		16±3	
F269W	i)12±3[#] ii)8±2 iii)8±2	9.5±2.3	i)<5[#] ii)<5	<5	i)42±5[#] ii)45±6	43.5±5.5	i)28±3[#] ii)22±2	25±2.5
wt control	10±3		<5		38±3		23±1	
F269Y	i)11±3[#] ii)8±2 iii)10±2	9.6±2.3	i)<5[#] ii)<5 iii)<5	<5	i)24±3[#] ii)20±3	22±3	i)15±2[#] ii)10±2	12.5±2
wt control	10±2		<5		25±3		14±2	
N266D	i)13±3[#] ii)10±3	11.5±3	i)<5[#] ii)<5	<5	i)52±5[#] ii)44±4	48±4.5	i)32±3[#] ii)32±4 iii)38±5	34.7±4
wt control	10±3[#]		<5[#]		50±5[#]		32±4[#]	
G177D	i)15±2[#] ii)12±3 iii)17±3 iv)20±3	16.3±2.8	i)21±3[#] ii)18±3 iii)15±2 iv)14±2	17.5±2.5	i)24±5[#] ii)28±4 iii)25±3	26.5±4.5	i)31±4[#] ii)27±3 iii)27±3	29.5±3.5
wt control	13±1		<5		36±4		20±2	

[a] K_d values are in nM and the error values (±) denote the range rather than the standard deviation.

[b] Powell *et al.*, 1993.

[#] indicates this result is shown in figure 3.8.

The mean K_d values for the mutant enzymes F269C, F269G, F269W, F269Y, and N266D, are in good agreement with the respective wild-type controls. However, the dissociation constants obtained using the non-specific DNA substrate (862/863) were consistently lower than those observed by Powell *et al.*, 1993. This may be the result of a slight difference in technique although it was not possible to determine the reason for the discrepancy. This highlights the importance of performing a wild-type control in such an experiment.

In the same manner as the wild-type, the F269C, F269G, F269W, F269Y, and N266D methylases display stronger binding for the specific substrate (860/409) than for the non-specific substrate (862/863) indicating that they are able to distinguish the target recognition site. This result indicates that the lack of catalytic activity associated with some of the mutants is probably not the result of a failure to interact with the DNA target site. All but one of the mutants bind both specific and non-specific DNA substrates more tightly in the presence of AdoMet indicating that not only are they are also able to bind AdoMet but that this interaction is productive in allowing stronger association with the DNA substrate.

In the same manner as the other mutant enzymes, the binding of G177D to the specific DNA is stronger than for the non-specific DNA although binding to the latter appears to result in complexes which are more prone to dissociation perhaps as a result of a decrease in stability and consequent dissociation within the gel. However, although G177D is able to distinguish specific from non-specific DNA substrate, importantly it does not appear to bind either substrate any differently in the presence of AdoMet. This may indicate that G177D does not bind AdoMet, or alternatively, on binding AdoMet the usual conformational changes which result in increased DNA binding are unable to take place.

3.8 AdoMet Binding

AdoMet binding was followed using three methods. The first method used the fluorescence of the extrinsic fluorophore ANS and was quantitative, providing a determination of the dissociation constant for AdoMet (section 3.7.1) The second method used gel filtration and provided a qualitative result (section 3.7.2). The third method used photo-crosslinking of AdoMet to methylase and provided an indirect, qualitative measure of interaction (section 3.7.3).

3.8.1 AdoMet Binding by Fluorescence of an extrinsic Fluorophore

The molecule ANS (1, 8-anilino-napthalene sulphonic acid) binds to hydrophobic patches and nucleotide-binding sites on proteins (Slavik, 1982; Chou *et al.*, 1989; Taylor *et al.*, 1992). The fluorescence of ANS in solution is very weak with an emission maximum at 515nm but on binding to proteins the fluorescence is enhanced and shifted to shorter wavelengths. The displacement of bound ANS by a ligand or substrate of the protein can be followed by the resulting decrease in fluorescence and dissociation constants can be calculated. This effect has been demonstrated for the wild-type *Eco* KI methylase and a value of 2-2.5μM obtained for the dissociation constant (K_d) of AdoMet (Powell *et al.*, 1993). Shi *et al.*, 1994, have suggested however that the displacement of ANS may be the result of preferential binding to an intermediate state. If ANS binds an intermediate state then the equilibrium can be shifted in favour of this state. The addition of ligands (e.g. AdoMet) which bind tightly to the native state would shift the equilibrium back to the native state and thus displace the bound ANS regardless of whether ANS binds in the nucleotide-binding site or not. However, in the case of the *Eco* KI methylase a second method for determining the dissociation constant of AdoMet (equilibrium dialysis) gave the same value indicating that ANS does not measurably shift the equilibrium in favour of an intermediate state and that therefore the use of the extrinsic fluorophore technique is valid in this instance (Powell *et al.*, 1993). The wild-type *Eco* KI methylase has been shown, by equilibrium dialysis, to have two AdoMet-binding sites which are either identical or so similar that there are no co-operative interactions between them (Powell *et al.*, 1993).

Dissociation constants for AdoMet were determined for all purified mutants apart from F269C which was slightly more prone to aggregation under the conditions utilised, preventing measurement of fluorescence during the titration. Fluorescence data were analysed using the Grafit data analysis programme (Leatherbarrow, 1990) with a ligand-binding equation for identical non-interacting sites. Figure 3.9 demonstrates the decrease in fluorescence observed when a 1μM methylase solution containing 50μM ANS is titrated with AdoMet. Table 3.9 details the dissociation constants (K_d values) for AdoMet obtained for each of the mutants.

Figure 3.9 (1 of 3)

AdoMet Binding by Methylase as determined by Titration of Methylase+ANS with AdoMet. The displacement of bound ANS and consequent reduction in fluorescence as a function of AdoMet concentration. The wild-type, F269G, F269W and F269Y methylases show a distinct reduction in fluorescence at an [AdoMet] of 1-8μM. G177D shows no displacement of ANS except at very high AdoMet concentrations.

Methylase concentration: 1μM
ANS concentration: 50μM

i) Wild-type
ii) F269G

Figure 3.9 (1 of 3)

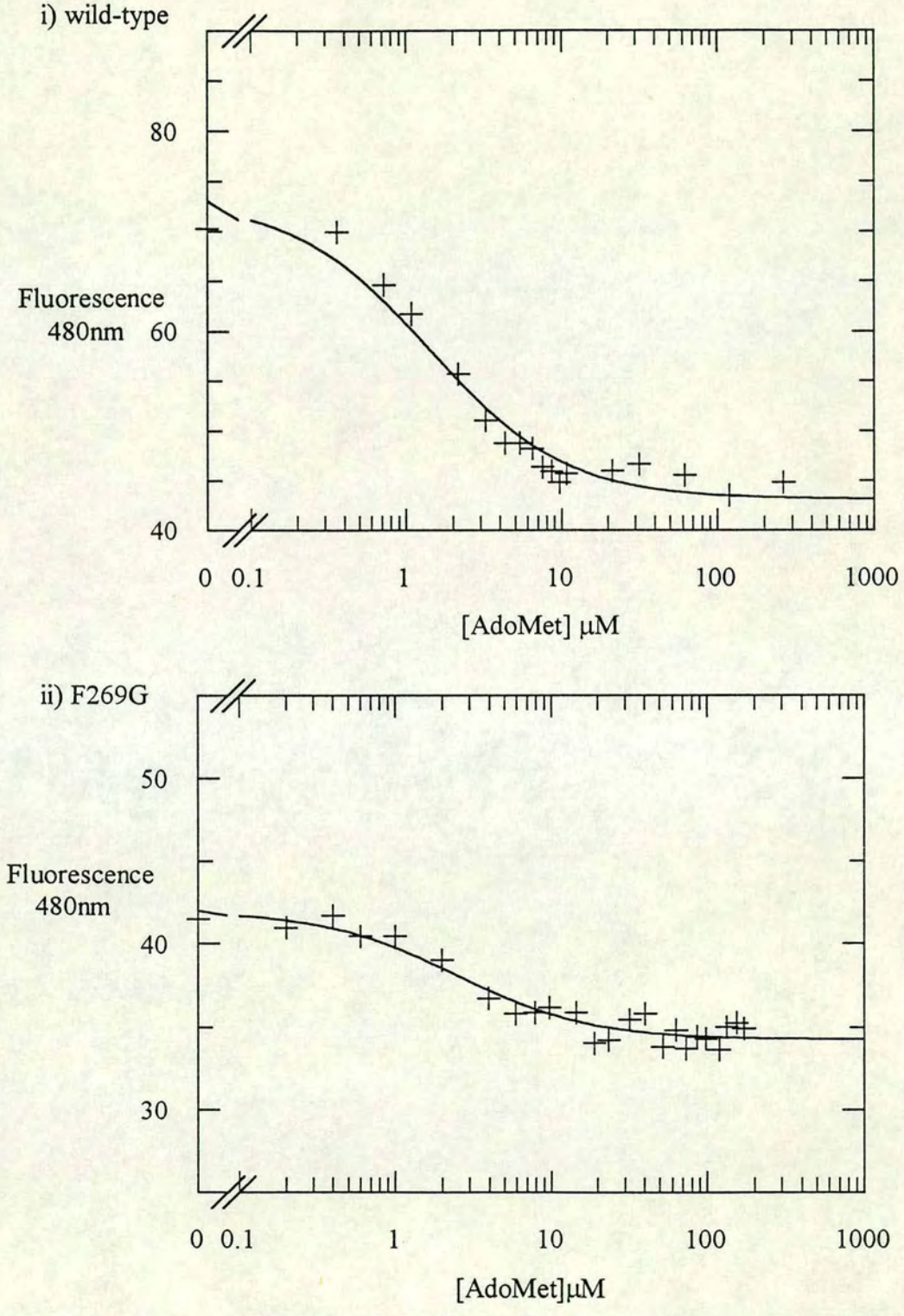
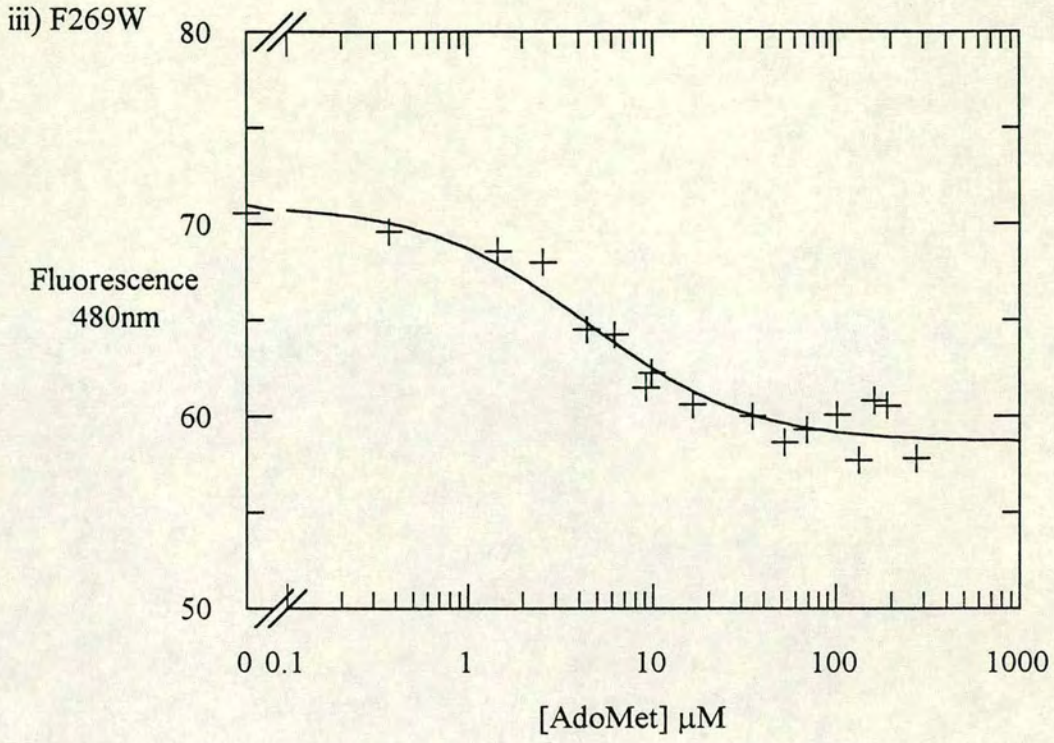


Figure 3.9 (2 of 3)

AdoMet Binding by Methylase as determined by Titration of Methylase+ANS with AdoMet.

i) F269W
ii) F269Y

Figure 3.9 (2 of 3)



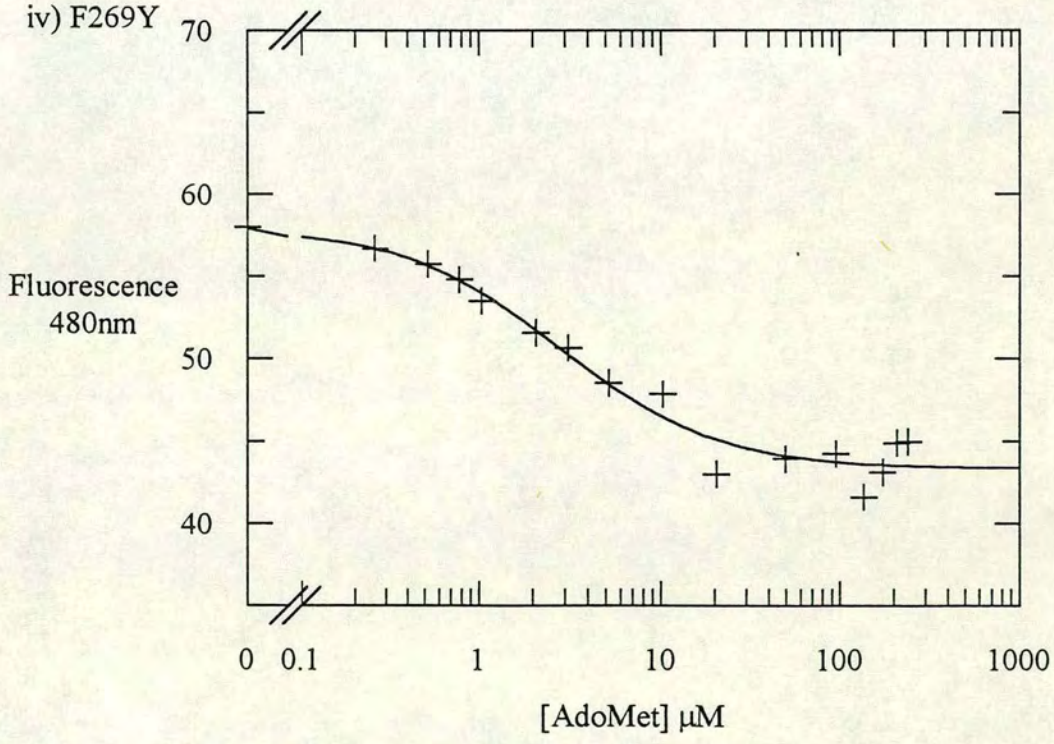


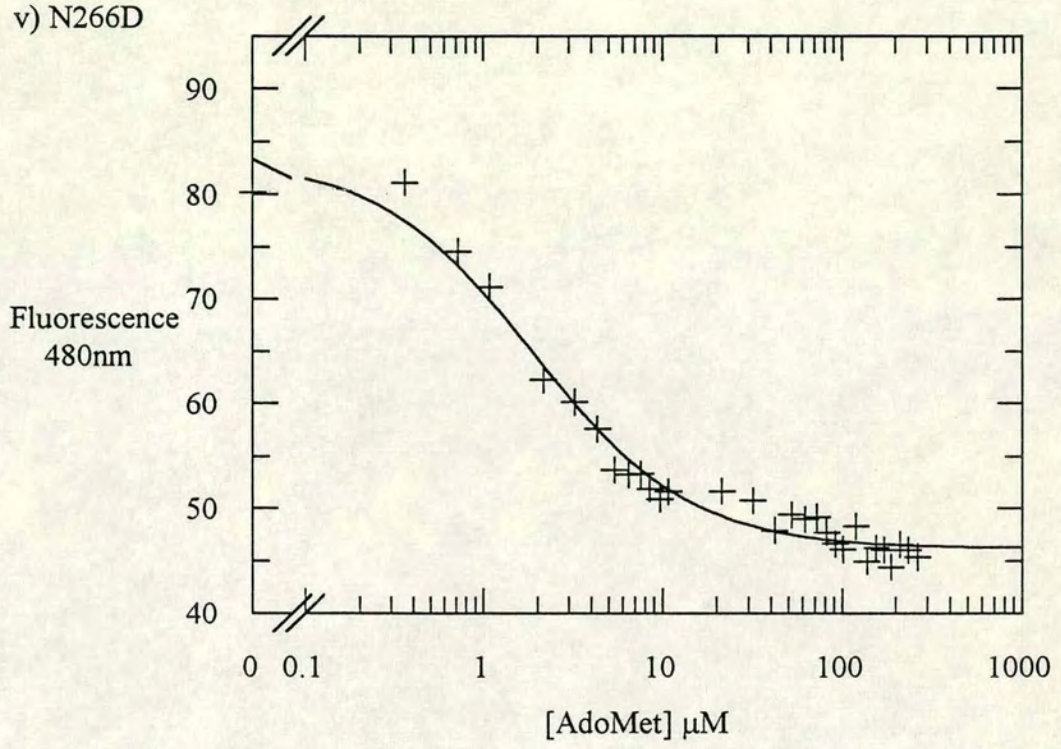
Figure 3.9 (3 of 3)

AdoMet Binding by Methylase as determined by Titration of Methylase+ANS with AdoMet.

i) N266D
ii) G177D

Figure 3.9 (3 of 3)

v) N266D



vi) G177D

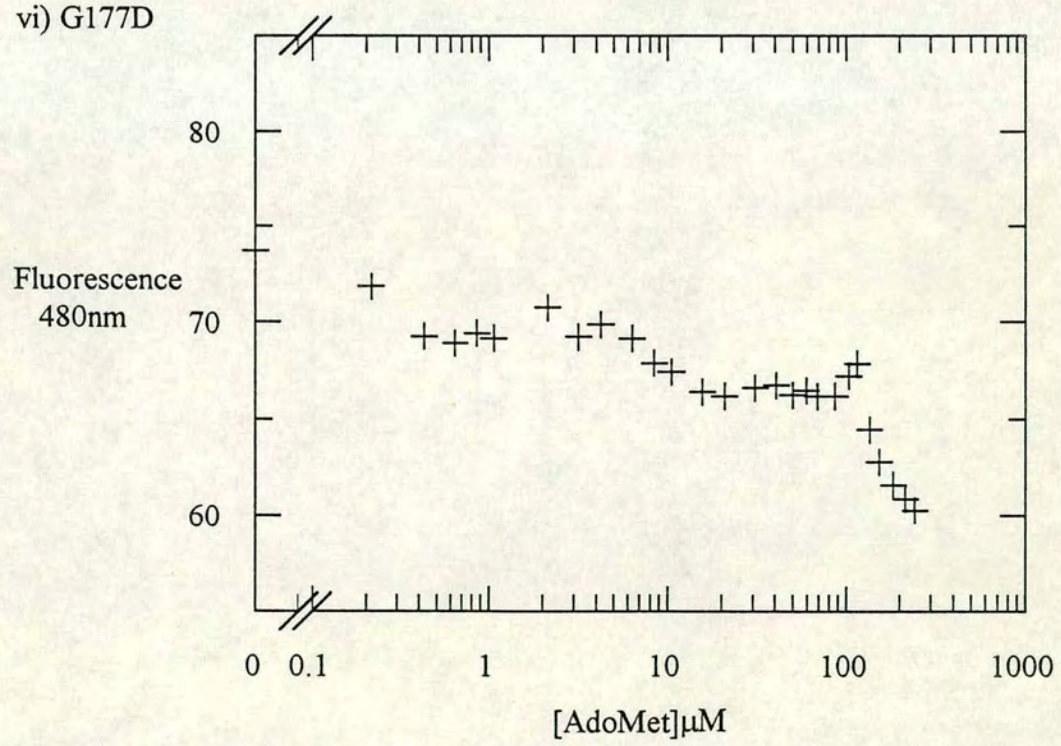


Table 3.9 Dissociation constants for AdoMet as determined by Titration of Methylase+ANS with AdoMet

Methylase	K_d (μM)	K_d (μM) Mean
wild-type		2.21±0.29#
G177D	No apparent K_d	
N266D	i) 1.9±0.2# iii)1.1±0 ii)1.2±0.2	1.6±0.2
F269C	ND	ND
F266G	i) 7.8±2.2 ii) 2.5±0.6 iii) 3.3±0.4 iv)3.8±0.43#	4.4±0.9
F269Y	i)1.5±0.8 ii)2.8±0.6# iii)1.3±0.9	1.9±0.8
F269W	i) 4.5±1.1# ii) 8.7±2.6	6.6±1.8

# indicates that this result is shown in figure 3.9
ND: not determined

The above results demonstrate that all but one of the mutant proteins bind AdoMet with approximately the same degree of interaction as the wild-type enzyme, their dissociation constants being in the range 2-7 μM. This result supports the DNA binding study which indicated that these mutant enzymes interact positively with AdoMet. However, G177D binds AdoMet very poorly, if at all (figure 3.9 vi). This

supports the DNA binding study which indicated that AdoMet was unable to elicit the increase in binding affinity associated with the wild-type enzyme.

It is not clear why F269C should have a tendency to aggregate under the conditions used in this experiment, particularly since denaturation by GuCl did not indicate any significant differences in stability when compared with the wild-type (section 3.6). The denaturation using GuCl was performed at pH 8.0 whereas the ANS titrations were performed at pH 6.5. It is possible that F269C is more unstable than the wild-type at pH 6.5 but not at pH 8.0 and hence differences in stability would only be obvious during experiments at the lower pH (the formation of a disulphide bond might affect the charge balance, leading to instability at a particular pH). Alternatively, F269C may display a local change in conformation which may enable ANS to bind more favourably to a less stable form hence causing aggregation. Alternative techniques were used to investigate the interaction of F269C with AdoMet (3.8.2 and 3.8.3).

3.8.2 AdoMet Binding by Gel Filtration

Gel filtration can be used to separate molecules of different molecular weight, for example, proteins and ligands. Ligand which is bound by protein will co-elute with the protein but free ligand will not. This technique has previously been used to characterise an *Eco* KI mutant deficient in AdoMet binding (Bühler and Yuan, 1978).

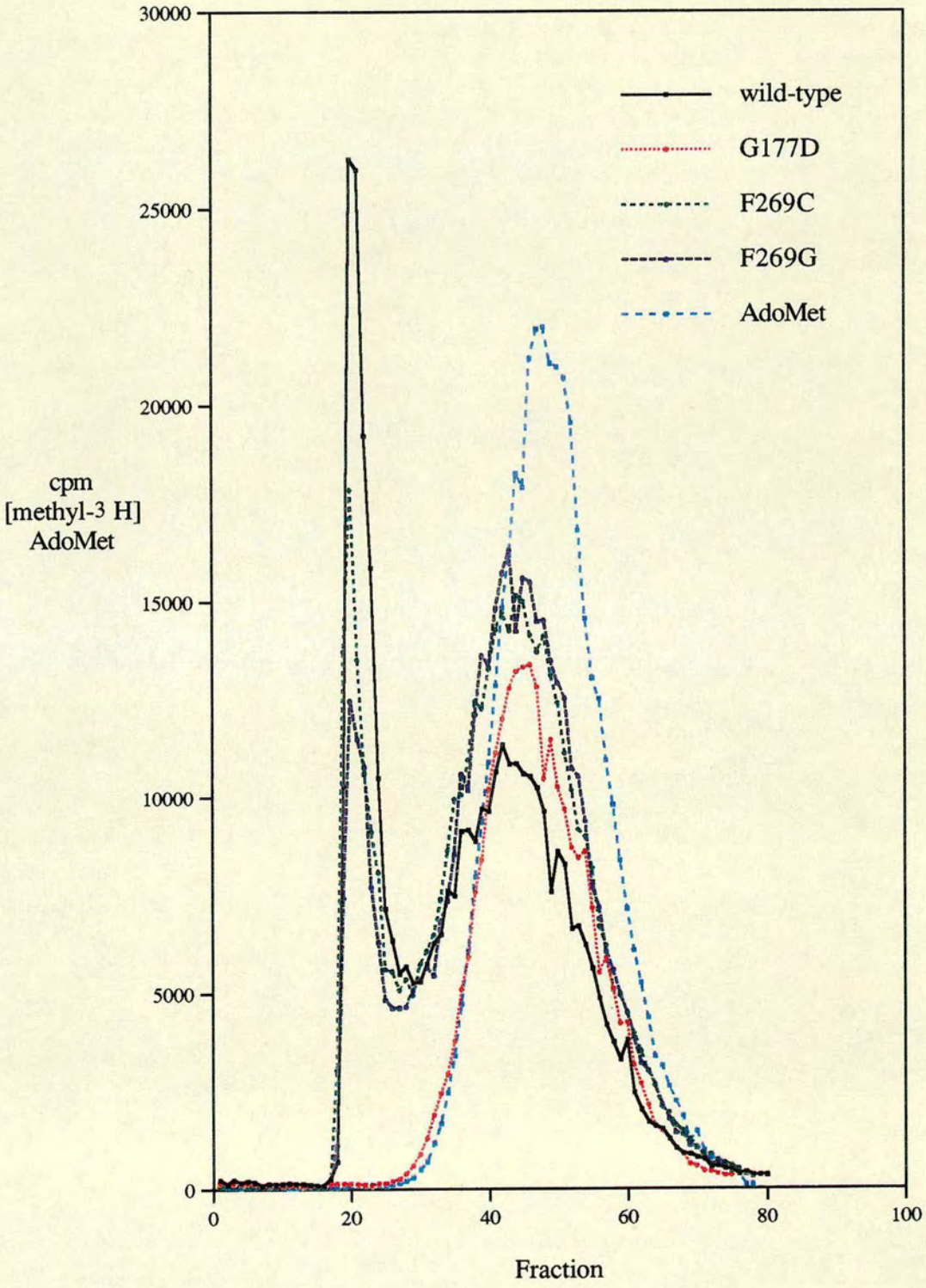
The following mutant methylases were investigated using this technique; F269C, F269G and G177D. F269C was chosen because this mutant was refractory to AdoMet binding analysis using the fluorescence technique as a result of a tendency to aggregate over a period of time under the conditions employed. F269G was chosen as an example of a catalytically inactive mutant enzyme which had indicated normal AdoMet binding using the fluorescence technique. G177D was chosen since the fluorescence technique (3.8.1) had indicated very poor binding to AdoMet.

Figure 3.10 is a graph of counts per minute (cpm [*methyl*-^{3}H]AdoMet) versus fraction number for the wild-type, F269C, F269G and G177D methylases. The wild-type plot shows a first peak at fraction 20. This corresponds to [*methyl*-^{3}H]AdoMet co-eluted with the protein. The second, broader peak at approximately fraction 45 corresponds to free [*methyl*-^{3}H]AdoMet. Similarly, both F269C and F269G also demonstrate a first peak at fraction 20 indicating that they too bind AdoMet. However G177D does not give a peak in the region of fraction 20 indicating that it does not bind AdoMet. Furthermore, BSA, a protein which does not bind AdoMet, failed to give a peak at fraction 20 when examined in this experiment (not shown). This result supports the fluorescence analysis of AdoMet binding, demonstrating that G177D binds very weakly, if at all, to AdoMet. Furthermore this technique demonstrates that F269C does bind AdoMet although it does not provide a value of the dissociation constant for the cofactor. (A modification of this technique where the gel is equilibrated with [*methyl*-^{3}H]AdoMet could be used to provide K_d values.)

Figure 3.10

AdoMet Binding by Wild-type, F269C, F269G, and G177D Methylases as determined by Gel Filtration. Elution of [*methyl* - ^{3}H]AdoMet with methylase on a Sephadex G25 gel filtration column. The peak at fraction 20 corresponds to protein-bound AdoMet whereas the peak at fraction 45 corresponds to free AdoMet. The wild-type, F269C and F269G methylases bind AdoMet but G177D does not.

Figure 3.10


30000
25000
20000
15000
10000
5000
0
cpm
[methyl-3 H]
AdoMet
wild-type
G177D
F269C
F269G
AdoMet
0
20
40
60
80
100
Fraction

3.8.3 AdoMet Binding by UV Cross Linking of Protein to [*methyl* - ^{3}H]AdoMet

The ability of a protein to undergo UV induced crosslinking to a ligand or substrate provides an indirect and qualitative description of the interactions taking place. The first use of tritium-labelled AdoMet as a photoaffinity label was for the phenylethanolamine *N* -methylase (Yu, 1983). Subsequently, UV induced [*methyl* - ^{3}H]AdoMet linkages have been demonstrated in several DNA and small molecule methylases (Som and Friedman, 1990; Reich and Everett, 1990; Wenzel and Guschlbauer, 1991; Subbaramaiah and Simms, 1992). The *Eco* KI wild-type methylase has been shown to form such linkages with [*methyl* - ^{3}H]AdoMet. The majority of the radioactivity (92%) was associated with HsdM and only 8% was associated with HsdS. The minor amount of crosslinking to HsdS is probably a result of the coincidental contact between AdoMet and HsdS when the cofactor is bound to HsdM (Powell *et al.*, 1993).

The mutant methylases fall into three classes with respect to their ability to undergo UV induced crosslinking to [*methyl* - ^{3}H]AdoMet. The first class are those mutants which display approximately the same level of crosslinking as the wild-type enzyme. This class comprises; N266D, F269C, F269G and F269Y. The second class are those mutants which have elevated levels of crosslinking. This class comprises only F269W. The third class are those mutants which have reduced levels and this class comprises only G177D. Figure 3.11 provides a visual comparison of the variation in UV induced crosslinking. Table 3.9 details the approximate levels of crosslinking for each of the mutants relative to the wild-type enzyme as determined by scanning of fluorographs.

Figure 3.11

AdoMet Binding by Wild-type, F269C, F269G, F269W, F269Y, N266D, and G177D Methylases as determined by UV Crosslinking of Protein to [*methyl*-^{3}H]AdoMet. G177D has a very low level of crosslinking and F269W has a very high level of crosslinking. F269Y, F269C, F269G and N266D have the same level of crosslinking as the wild-type methylase.

Figure 3.11

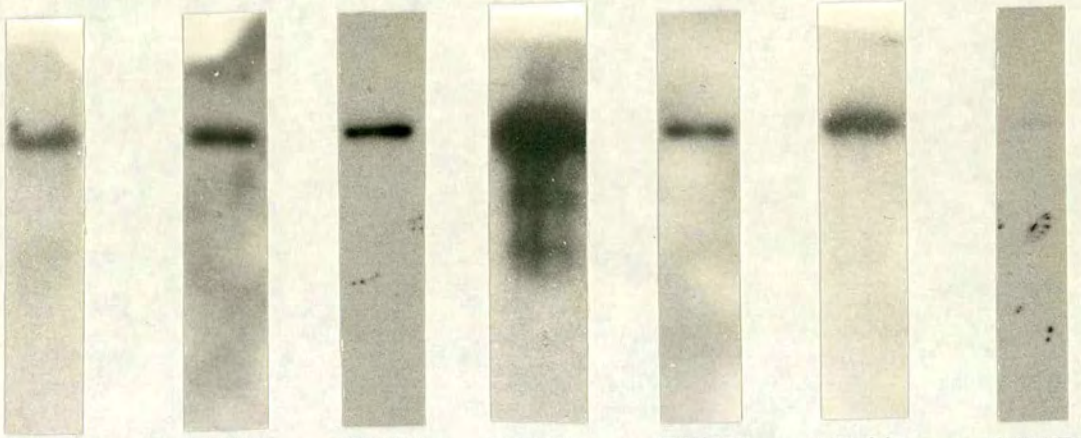


Table 3.9 Levels of UV induced Crosslinking between [*methyl* - ^{3}H]AdoMet and Methylase Mutants, relative to Wild-type.

Methylase	UV induced crosslinking to [*methyl* - ^{3}H]AdoMet (Relative to wild-type)
wild-type	1
G177D	0.02-0.04
N266D	1
F269C	1
F269G	1
F269Y	1
F269W	40-50

The considerably decreased level of UV induced [*methyl* - ^{3}H]AdoMet crosslinking of the G177D methylase relative to the wild-type may indicate that the interaction of this mutant with AdoMet is very weak. Alternatively, the photoreactive amino acid residue(s) responsible for crosslinking in the wild-type enzyme may no longer, in G177D, be in a position favourable for crosslink formation to occur. It is unlikely that Gly 177 itself forms UV induced crosslinks to AdoMet since glycine is not known to be a photoreactive residue. A more likely explanation is that substitution of Gly in G177D has caused a conformational change in the protein resulting in a reduction of binding affinity for AdoMet and hence a lowered level of UV induced crosslinking. The conformational change may be one in which the overall structure of the protein has been altered or it may be such that only the local conformation of the AdoMet-binding region is affected. The latter alternative would seem more likely given that the conformational stability of G177D appears to be the same as the wild-type enzyme (as indicated by GuCl unfolding) and that this mutant enzyme, at 25°C, purifies in the same manner as the wild-type enzyme. This latter alternative would infer that Gly 177 and therefore region I, constitutes part of the AdoMet-binding region, a hypothesis consistent with the two other methods used to analyse AdoMet binding. This is also consistent with the inability of AdoMet to effect enhanced DNA binding by the G177D methylase.

The greatly increased level of crosslinking in F269W may indicate an increase in binding affinity for AdoMet. Alternatively, the substitution of Trp for Phe 269 has caused a conformational change which has resulted in the residue(s) which crosslink to AdoMet being placed in a more favourable position.

Tryptophan, however, is well documented as being the most photoreactive amino acid (Asquith and Rivett, 1971). The photoionisation of tryptophan (and to a much lesser extent tyrosine) constitute a major initial photochemical reaction in the photolysis of proteins. Ovine trypsin, for example, is inactivated by photolysis of a tryptophan residue adjacent to the catalytic serine (Grossweiner *et al.*, 1976).

Therefore, given the relatively high photoreactive nature of tryptophan, it would seem most likely that Trp 269 itself forms a covalent adduct with AdoMet resulting in an elevated level of crosslinking. This hypothesis would suggest that this tryptophan residue, and therefore region II, is in close proximity to bound AdoMet.

In conclusion the UV crosslinking of methylase mutants to [*methyl*-^{3}H]AdoMet indicates that elements from both regions I and II contribute to the AdoMet-binding site in the *Eco* KI methylase.

Chapter Four

DISCUSSION

The *hsdM* gene of the type I methylase *Eco* KI encodes amino acid sequence motifs common to both N6-adenine methylases and methylases in general (section 1.4).

The first of these motifs, region I (F X G X G), is common not only to N6-adenine methylases but to all methylases that utilise AdoMet as the methyl donor (Ingrosso *et al.*, 1989). The universal occurrence of this motif in such enzymes suggests a key role in the interaction with AdoMet (section 1.4). Although the motif is not well conserved in M.*Eco* KI (tables 1.2 and 1.3) the invariant Gly residue (G) is present and it is probable that this residue fulfils a similar role in all methylases containing the motif.

The second of these motifs, region II (S/N/D P P F/Y), is found in both N6-adenine methylases and in N4-cytosine methylases, both of which catalyse the methylation of an extracyclic amino group. In N6-adenine methylases the region II motif is characterised by two invariant proline residues bordered on the amino side by an asparagine or aspartate residue, and on the carboxyl side by a phenylalanine or tyrosine residue. In N4-cytosine methylases a serine residue replaces the asparagine/aspartate residue (figure 1.2). Suggestions for the role of region II have included; part of the active site for adenine methylation (Loenen *et al.*, 1987), adenine recognition (Chandrasegaran and Smith, 1988), and AdoMet binding (Lauster *et al.*, 1987; Guschlbauer, 1988) (section 1.4). The motif displays a degree of similarity to the characterised P C motif found in C5-cytosine methylases and therefore may fulfil a similar function, i.e. a central role in catalysis (Klimasauskas *et al.*, 1989).

The results described in this thesis allow several conclusions to be made regarding the roles of the region I and region II motifs as found in the *Eco* KI methylase and, by extrapolation, the roles played in other methylases which share these motifs.

Mutations made in region II indicate that the consensus sequence for this region has a degree of stringency since amino acid substitutions which remain within the consensus do not necessarily retain activity. The single mutation N266D and the double mutation N266D/F269Y confer phenotypes indicative of no *in vivo* methylase activity (section 3.3). Furthermore, assay of the N266D mutant enzyme

indicates no *in vitro* activity but its DNA binding, AdoMet binding, and stability, are in all respects identical to the wild-type. The third mutation which remains in the overall consensus, F269Y, is identical to the wild-type in all respects except for a reduction in the absolute level of *in vitro* methylase activity (section 3.5). Although a biochemical analysis of the double mutant, N266D/F269Y was not performed, it would seem likely that the reason(s) for its inactivity are the same as for the single mutant, N266D, because the F269Y mutation alone confers catalytic activity.

Three mutations were made in region II which were outside the consensus sequence: F269C, F269G, and, F269W. Both the F269C and F269G mutations confer phenotypes indicative of no *in vivo* methylase activity. In support of this result, enzyme assays indicate no *in vitro* activity for these methylases. In all other respects, F269G methylase is identical to the wild-type enzyme. F269C methylase is also identical to the wild-type enzyme apart from a tendency to show more aggregation under certain conditions. The tendency to aggregate may be a pH effect, or perhaps a local change in conformation (section 3.8.1). The third mutation outside the consensus sequence, F269W, confers a phenotype indicative of an intermediate level of *in vivo* methylase activity. In agreement with this intermediate level, the F269W methylase has a considerably reduced, but detectable, level of *in vitro* activity (section 3.5). The dissociation constants of the F269W methylase for DNA and AdoMet are the same as those of the wild-type indicating that the affinity of this mutant for the two substrates is not reduced. However, the level of crosslinking of this mutant to [*methyl* - ^{3}H]AdoMet is greatly elevated relative to the wild-type enzyme (40-50x) indicating, significantly, that the Trp residue is likely to be in close proximity to bound AdoMet (3.8.3). Interestingly, the nucleotide sequence of the N6-adenine methylase, *Vsp* I, has been found to encode a region II motif of the sequence N P P W (Degtaryev *et al.*, 1993). This further indicates that the presence of an aromatic residue is required at the last position of the region II motif and that the overall consensus sequence for this motif can be extended to S/N/D P P F/Y/W. Given the enhanced UV crosslinking to [*methyl* - ^{3}H]AdoMet of the M.*Eco* KI F269W mutant enzyme one might predict that the M.*Vsp* I wild-type would also display a high level of crosslinking. Mutagenesis of this Trp residue within M.*Vsp* I would be expected to result in decreased levels of crosslinking to [*methyl* - ^{3}H]AdoMet.

A comparison can be made between the region II mutagenesis study described in this thesis and three other investigations of the motif, in the N6-adenine methylases; M.*Fok* I, *Eco* Dam, and T4 Dam.

In a successful attempt to inactivate the N6-adenine methylase M.*Fok* I the two copies of the region II motif were changed from D P P Y to A P P Y and G P P Y indicating the importance of the first residue of this motif (section 1.4, Sugisaki *et al.*, 1989). However, since no biochemical analysis was performed it was not determined why these amino acid changes had resulted in inactivity.

In a site-directed mutagenesis study of the region II motif of T4 Dam (D P P Y), substitution of the first Pro residue for Ala, Thr, Arg or Cys resulted in enzymes with (respectively) 8%, 1.5%, 1% and <1% of wild-type activity in cell extracts. The Pro→ Ala and Pro→ Gly mutant enzymes were purified and found to have reduced affinity for AdoMet but unchanged affinity for DNA. It was suggested therefore that region II contributes to the binding site for AdoMet (Kossykh *et al.*, 1993).

The closely related N6-adenine methylase *Eco* Dam also has the region II motif D P P Y. In a mutagenesis study seven amino acid substitutions of the region II Pro residues all resulted in enzyme inactivity. In addition all of these mutants had either undetectable or low levels of UV crosslinking to [*methyl*-^{3}H]AdoMet. Furthermore, substitution of Asp for Ser, Asn, or Gly resulted in inactivity and undetectable crosslinking to [*methyl*-^{3}H]AdoMet. DNA binding was analysed for only one mutant protein (Asp→ Asn) and was also found to be affected. The *Eco* Dam wild-type binds specific DNA more tightly than non-specific DNA (K_d values 300nM and 3μM, respectively) and binding to both substrates is enhanced by AdoMet (K_d values 60nM and 1.2μM, respectively) (Bergerat and Guschlbauer, 1990). However, the Asp→ Asn mutant bound both substrates with the same K_d (20nM) and in the presence of AdoMet, binding was unchanged (Guyot *et al.*, 1993). Hence this mutation was able to disrupt both AdoMet binding and confer abnormally tight binding of DNA.

The mutagenesis studies of both T4 Dam and *Eco* Dam therefore suggest that in these enzymes the region II motif is in close proximity to the binding site for AdoMet (or is part of the AdoMet-binding site) and, at least in *Eco* Dam, is also involved in DNA interaction. In contrast, the M.*Eco* KI region II mutations made in this thesis (F269C, F269G, N266D) result in catalytic inactivity but do not affect the binding of either AdoMet or DNA. The region II studies on T4 Dam and *Eco* Dam (Kossykh *et al.*, 1993; Guyot *et al.*, 1993) used only one method, protein

crosslinking to [*methyl* - ^{3}H]AdoMet, as a measure of AdoMet binding. It is possible that these mutants still bind AdoMet but the photoreactive residue(s) are no longer appropriately positioned for either covalent crosslinking or, in the case of *Eco* Dam Asp→Asn, for correct DNA interaction to occur. Alternatively, and more likely, given the number of amino acid changes made, the mutations in T4 Dam and *Eco* Dam may indeed have resulted in a failure to bind AdoMet. In agreement with the results in this thesis, the *Eco* Dam D→ N and D→ S mutations demonstrate that changes within the overall consensus of N6/N4 DNA methylases do not necessarily retain activity.

As discussed above, the effects of amino acid changes within the region II motif can include: (i) alteration to the binding of AdoMet and DNA (Kossykh *et al.*, 1993; Guyot *et al.*, 1993), (ii) disruption of catalysis without affecting binding of either substrate (this thesis). This combination suggests that this motif is in close proximity to the AdoMet-binding site and also interacts with DNA, perhaps as a consequence of an involvement in catalysis.

A further indication that there is interaction between region II and DNA is provided by an observation made on the phage P22 Mnt repressor (section 1.4, Chandrasegaran and Smith, 1988). An inactive mutant repressor containing the sequence D P P F (i.e the amino acid sequence found in region II) is suppressed by an operator mutation containing the modified base, 6-methyladenine. The unmodified adenine base, however, is unable to suppress the mutation, implying specific interaction between the amino acid sequence D P P F and 6-methyladenine. Similarity has been identified between the region II motif of N6-adenine and N4-cytosine methylases, and the P C motif of C5-cytosine methylases (table 1.2, Klimasauskas *et al.*, 1989). The Cys residue of the P C motif has been demonstrated to be the catalytic nucleophile in such methylases (Chen *et al.*, 1991). Mutation of the Cys residue to Gly in M.*Hha* I resulted in inactivity as a result of abnormally tight DNA binding (Mi and Roberts, 1993). Abnormally tight binding of DNA was also the result of the region II mutation D→N in the N6-adenine methylase *Eco* Dam (Guyot *et al.*, 1993). This finding, taken into account with the sequence similarity observed between the region II motif and the P C motif, may suggest analogous roles for these regions, in which case region II would be involved in catalysis.

This thesis describes one mutation, G177D, made in the region I motif of M.*Eco* KI. As well as conferring a phenotype indicative of no *in vivo* methylase activity, the mutant enzyme has no *in vitro* activity (sections 3.3 and 3.5). Unlike the inactive mutants made in region II, the G177D enzyme differs from the wild-type with regard to the binding of AdoMet. All three methods used to investigate AdoMet binding indicate that the G177D enzyme does not bind the cofactor (sections 3.8.1, 3.8.2, 3.8.3). Predictably, the failure to bind AdoMet results in K_d values for DNA which are unaffected by the presence of AdoMet (section 3.7). However, the G177D enzyme is able to differentiate between specific and non-specific DNA, indicating that sequence specificity is unchanged by the mutation. Another distinguishing feature of G177D is its insolubility when expressed at 37°C. Fortunately, when expressed at 25°C, a substantial proportion of the protein is soluble thus enabling its purification (section 3.4). However, the reason(s) for the insolubility of the protein at 37°C is not apparent at 25°C when investigated by GuCl denaturation (section 3.6). Therefore the lack of *in vitro* activity at the lower temperature and the failure to bind AdoMet do not appear to be the result of a reduction in stability. It is well documented that protein solubility can be greatly changed by single amino acid substitutions even if no conformational change occurs. In many instances the reason for the insolubility of a protein is not necessarily the final conformation of the protein (Thatcher and Hitchcock, 1994). In a mutagenesis study of phage tail spike protein, for example, some amino acid substitutions were found to induce insolubility when expressed at 37°C but at lower temperatures were found to be soluble. *In vitro* these mutants were as stable as the wild-type and hence the mutations were interpreted as causing lesions in the folding pathway rather than causing any change to the conformation of the fully folded protein (Mitraki and King, 1992). This is also a valid interpretation of the G177D mutant enzyme.

Two other mutations in region I of M.*Eco* KI have been documented. A temperature sensitive mutation in *hsdM* has been described which encodes an enzyme defective in AdoMet binding (Hubacek, 1973; Bühler and Yuan, 1978). Upon sequencing, the mutation was found to comprise three amino acid changes; P174S, T280S and G388D (Daniel, A., unpublished). The first of these three changes is in the region I motif (table 1.2). Site-directed mutagenesis was used to create the single mutation, P174S, and the conferred phenotype was temperature sensitive, consistent with this amino acid change being responsible for the phenotype of the triple mutant noted by Bühler and Yuan. A second mutation in region I has been described, D173G, the phenotype of which is indicative of no *in vivo*

methylation activity (Kelleher, 1990). When overexpressed, this mutant is insoluble at both 25°C and 37°C. Therefore it appears that mutations made in region I of M.*Eco* KI have a distinct tendency to induce insolubility as well as interfering with the binding of AdoMet.

Mutations have been made in the region I motif of two other enzymes. In the multispecific phage C5-cytosine methylase SPR, which methylates three different DNA target sequences, mutation of the highly conserved Gly residue to Glu resulted in a loss of capacity to methylate all three of the sequences normally methylated by the wild-type (Wilke *et al.*, 1988). However the nature of the mutant enzymes was not investigated. Preliminary results of mutations (unspecified) made in the region I motif of the N6-adenine methylase T4 Dam and of the C5-cytosine methylase *Eco* RII, indicate that some mutations may result in changes to catalytic activity without affecting the affinity for either AdoMet or DNA (Kossykh *et al.*, 1994). This is the same phenotype caused by the region II mutations, F269C, F269G, and N266D, described in this thesis. The T4 Dam and M.*Eco* RII mutant enzymes were interpreted as perhaps having a defect in the positioning of AdoMet relative to the catalytic centre.

Hence mutations made in region I do not appear to affect DNA binding but they can cause: (i) loss of AdoMet binding and temperature sensitivity (this thesis), (ii) loss of catalytic activity without loss of AdoMet binding, possibly as a result of an alteration in the positioning of AdoMet with respect to the catalytic centre.

The combined results of mutagenesis investigations in regions I and II suggest that: (i) both regions constitute part of the AdoMet-binding site or at least are in close proximity to bound AdoMet (and therefore in close proximity to each other), (ii) region II is involved in DNA interaction perhaps as a contribution to catalysis.

The determination of the crystal structure of the C5-cytosine DNA methylase *Hha* I has provided intriguing insights into the mechanism used by this class of enzymes (section 1.5). In the crystal structure of M.*Hha* I complexed with the cofactor AdoMet, the region I motif forms a tight turn between a β strand and an α helix (Cheng *et al.*, 1993). The invariant glycine residue within the turn may well be essential for its exact conformation since the lack of a side chain in glycine enables

conformations which are usually prohibited (Branden and Tooze, 1991). In the M.*Eco* KI G177D mutant this turn is unlikely to be correctly formed and this would account for the failure to bind AdoMet. The AdoMet-binding structure of M.*Hha* I resembles the β–α–β nucleotide-binding motif, or Rossmann fold structure (Rossmann *et al.*, 1974). The lack of obvious sequence similarity between the AdoMet-binding region of M.*Hha* I and the nucleotide-binding motif is typical of proteins with a Rossmann fold structure. It appears that in such proteins it is the structural framework which is preserved rather than specific amino acid contacts (Branden and Tooze, 1991).

A second methylase structure has recently been determined, that of catechol O-methylase (COMT) (Vidgren *et al.*, 1994). Although this enzyme is not a DNA methylase, the elements of its AdoMet-binding region are similar to those of M.*Hha* I, comprising a β–α–β structure and resembling the Rossmann fold. As in M.*Hha* I, the turn contains a glycine residue (table 1.3).

Very recently, the first structure of an N6-adenine methylase has been determined. The structure is of the *Taq* I methylase in a binary complex with AdoMet (Labahn *et al.*, 1994). The enzyme has a very similar overall structure to that of M.*Hha* I, comprising two domains between which is a cleft large enough to accommodate a DNA substrate. The conserved motifs of regions I and II are both located in the N-terminal domain and are of the sequence A C A H G and N P P Y respectively (table 1.2). Notably, region I is poorly conserved and the highly conserved Gly residue is replaced by Ala, although one glycine remains. Despite this lack of strong similarity to the region I consensus, the β–α–β structure seen in M.*Hha* I and COMT is still observed. Elucidation of this structure validates one of the conclusions drawn from the mutagenesis studies in regions I and II. As hypothesised, the two motifs in these regions do contribute to the binding of AdoMet. Together the motifs form the walls of a cavity which constitutes a pocket in which AdoMet is bound. Both motifs have sharp turns. In region I the turn appears to be mediated by a glycine residue (G51) and similarly in region II, the two proline residues (P106, P107) appear to mediate the turn. Unfortunately the structure does not include a DNA substrate and therefore these contacts are not known.

All three of the methylase structures display similar strategies for the binding of AdoMet and this suggests that all methylases utilising AdoMet might possess the same elements of secondary structure in the formation of a binding pocket. All three enzymes interact with the following AdoMet moieties: the N6 of adenine, the O2 of ribose, and the methionine amino and carboxyl groups. In M.*Taq* I and COMT there are van der Waals contacts to the adenine but these are replaced by hydrophobic interactions in M.*Hha* I. Further interactions are made in M.*Hha* I to the adenine N1 and N3 positions. In COMT there are further interactions to the adenine N1 and ribose O1 positions. Although the positions on the AdoMet molecule with which the enzymes interact are largely the same, the amino acids responsible are not necessarily the same. For example, the interaction (H-bond) to the methionine carboxylate is made by glutamate and threonine residues in M.*Taq* I, by a glutamine residue in M.*Hha* I, and by a backbone nitrogen in COMT.

In M.*Taq* I all but one of the amino acids interacting with AdoMet are highly conserved in the γ class of N6-adenine and N4-cytosine methylases (section 1.4, Winter, M., personal communication) . It is very likely therefore that the structure of the AdoMet-binding pocket and the amino acid contacts made by these enzymes, which include M.*Eco* KI, are the same as those displayed by M.*Taq* I. However, preliminary sequence alignments between the three classes, γ , α, and β, indicates that the latter two classes probably use different residues to interact with AdoMet.

In the M.*Taq* I-AdoMet structure, there is a distance of 15Å between the methyl donor, AdoMet, and the cleft which would accommodate the DNA substrate. This distance would be too large to enable the transfer of a methyl group onto DNA. The M.*Hha* I-AdoMet structure presented a similar anomaly when it was determined. However, solving of the tertiary complex M.*Hha* I-AdoHcy-DNA accounted for this large spatial separation (section 1.5). Upon binding AdoMet, the active site P C motif of M.*Hha* I moves nearly 25Å, taking it into the DNA minor groove. In addition, the cytosine substrate is completely rotated out of the DNA helix and into the catalytic site of the enzyme (figure 1.5). This allows the enzyme to perform catalysis at the C5 position of cytosine, a site which is relatively inaccessible when cytosine is in its normal base-paired conformation. Similarly, in base-paired DNA, the adenine N6 and the cytosine N4 positions are also inaccessible to an enzyme and it seems likely therefore that a similar base-flipping mechanism occurs in methylases of these classes. If base-flipping is a feature of N6-adenine and

N4-cytosine methylases, it would account for the region II mutagenesis studies in N6-adenine methylases which have indicated that there is interaction between the region II motif and DNA. It would also account for the the observation made on the phage P22 Mnt repressor which implied specific interaction between the amino acid sequence D P P F and 6-methyladenine (this chapter). A base-flip in the N6-adenine methylases might position the target adenine close to the region II motif and following catalysis would result in the juxtaposition of the motif with N6-methyladenine.

Base-flipping has been postulated in another DNA-modifying enzyme for which a crystal structure has recently become available. The structure of the T4 enzyme β-glucosyltransferase (T4 BGT) has been determined in the presence and absence of the its substrate uridine diphosphoglucose (UDPG) (Vrielink *et al.*, 1994). The enzyme transfers this molecule onto the modified base 5-hydroxymethylcytosine. Like the *Hha* I and *T aq* I methylases, the structure possesses a Rossmann nucleotide-binding motif despite there being no similarity between the amino acid sequences of T4 BGT and the methylases. T4 BGT has two domains which surround a cleft which is large enough to accommodate a DNA substrate. A channel extends from the cleft into the enzyme and it is here that the substrate UDPG is bound. The spatial separation between the bound UDPG and the cleft suggests that in order for transfer to occur, a large conformational change is necessary in either the enzyme or the DNA, or perhaps in both, as is the case in the mechanism of M.*Hha* I. A base-flip in which 5-hydroxymethylcytosine swings out of the DNA helix and into the cleft containing UDPG could contribute to this change. The structure of the C-terminal domain of another enzyme, the DNA repair protein Ada O^6-methylguanine-DNA methyltransferase, has recently been determined (Moore *et al.*, 1994). This enzyme belongs to a family of proteins (O^6-methylguanine methylases) which have the P C motif as found in C5-cytosine methylases (section 1.4, Rydberg, *et al.*, 1990). In the human enzyme the Cys residue has been shown to be the acceptor for the excised methyl group (von-Wronski *et al.*, 1991). The structure indicates that the P C motif in Ada is buried within the enzyme and a conformational change would be required in order to effect catalysis. However, the structure does not include the DNA substrate and it is not known therefore whether the mechanism involves the extraction of a base as seen in the structure of M.*Hha* I-AdoHcy-DNA, but this remains a possibility.

In conclusion this thesis demonstrates that elements from both regions I and II contribute to the AdoMet-binding site in the *Eco* KI methylase (Willcock *et al.*, 1994). Furthermore, elements from region II may interact with DNA by means of an involvement in catalysis. By extrapolation it seems likely that the same will be true for all N6-adenine and N4-cytosine methylases since they all share the respective motifs found in these regions. These findings are confirmed by the recent determination of the structure of the N6-adenine methylase *Taq* I complexed with AdoMet.

REFERENCES

Abadjieva, A., J. Patel, M. Webb, V. Zinkevich, and K. Firman. 1993. Nucl.Acid.Res. **21**: 4435-4443.

Adams, R.L.P., and R.H. Burdon. 1985. DNA structure and the effect of methylation. Chapter 3 in **The molecular Biology of DNA Methylation** (Springer Series in Molecular Biology). New York Inc.

Appleyard, R.K. 1954. Genetics **39**: 440-452.

Arber, W. 1964. Virology **23**: 173-182.

Arber, W. 1965. J.Mol.Biol. **11**: 247-256.

Arber, W., and D. Dussoix. 1962. J.Mol.Biol. **5**: 18-36.

Arber, W., and D. Wauters-Willems. 1970. Mol.Gen.Genet. **108**: 203-217.

Asquith, R.S., and D.E. Rivett. 1971. Biochim.Biophys.Acta. **252**: 111-116.

Benton, W.D., and R.W. Davies. 1977. Science **196**: 180-182.

Bergerat, A.,and W. Guschlbauer. 1990. Nucl.Acids.Res. **18**: 4369-4375.

Bertani, G., and J.J. Weigle. 1953. J.Bacteriol. **65**: 113-121.

Bestor, T.H. and V.M. Ingram. 1983. Proc.Nat.Acad.Sci. USA **80**: 5559-5563.

Bickle, T.A., C. Brack, and R. Yuan. 1978. Proc.Roy.Acad.Sci. **75**: 3099-3103.

Bickle, T.A., and D.H. Krüger. 1993. Microbiol.Rev. **57**: 434-450.

Biggin, M.D., T.J. Gibson, and G.F. Hong. 1983. Proc.Nat.Acad.Sci. USA **80**: 3963-3965.

Birnboim, H.C., and J. Doly. 1979. Nucl.Acids.Res. **7**: 1515-1523.

Boyer, H.W. 1971. Ann.Rev.Microbiol. **25**: 153-176.

Boyer, H.W., and D. Roulland-Dussoix. 1969. J.Mol.Biol. **41**: 459-472.

Branden, C., and J. Tooze. 1991. In **Introduction to Protein Structure** Garland Press.

Brooks, J.E., J.S. Benner, D.F. Heiter, K.S. Silber, L.A. Sznyter, T. Jager-Quinton, L.S. Moran, B.E. Slatko. G.G. Wilson, and, D.O. Nwankwo. 1989. Nucl.Acids.Res. **17**: 979-997.

Bühler, R., and R. Yuan. 1978. J.Biol.Chem. **253**: 6756-6760.

Burckhardt, J., J. Weisemann, D.L. Hamilton and R. Yuan. 1981. J.Mol.Biol. **153**: 425-440.

Chandrasegaran, S., and H.O. Smith. 1988. In **Structure and Expression** (M.H. Sarma, and R.H. Sarma, eds.). Adenine Press, Guilderland, NY. Vol. 1, pp. 149-196.

Chen, L., A.M. MacMillan, W. Chang, K. Ezaz-Nikpay, W.S. Lane, and G.L. Verdine. 1991. Biochemistry **30**: 11018-11025.

Cheng, S.-C., G. Herman, and P. Modrich. 1985. J.Biol.Chem. **260**: 191-194.

Cheng, X., S. Kumar, J. Posfai, J.W. Pflugrath, and R. Roberts. 1993. Cell **74**: 299-307.

Chou, W.-Y., C. Bieber, and K.S. Matthews. 1989. J.Biol.Chem. **264**: 18309-18313.

Cooper, L.P., and D.T.F. Dryden. 1994. J.Mol.Biol. **236**: 1011-1021.

Cowan, G.M., A.A.F. Gann, and N.E. Murray. 1989. Cell **56**: 103-109.

Daniel, A.S., F.V. Fuller-Pace, D.M. Legge, and N.E. Murray. 1988. J.Bacteriol. **170**: 1775-1782.

Degtaryev, S.K., E.A. Prikhod'ko, G.G. Prikhod'ko, and V.N. Krasnykh. 1993. Nucl.Acids.Res. **21**: 2015.

Dev, I.K., B.B. Yates, J. Leong, and W.S. Dallas. 1988. Proc.Natl.Acad.Sci.USA. **85**: 1472-1476.

Dryden, D.T.F., L.P. Cooper, and N.E. Murray. 1993. J.Biol.Chem. **268**: 13228-13236.

Dubey, A.K., and R.J. Roberts. 1992. Nucl.Acids.Res. **20**: 3167-3173.

Endlich, B., and S. Linn. 1985. J.Biol.Chem. **260**: 5720-5728.

Ford, K., C. Taylor, B. Connolly, and D.P. Hornby. 1993. J.Mol.Biol. **230**: 779-786.

Fratini, A.V., M. Kopka, H.R. Drew, and R.E. Dickerson. 1982. J.Biol.Chem. **257**: 14686-14707.

Frederick, C.A., G.J. Quigley, G.A. van der Marel, J.H. van Boom, A.H.-J.Wang, and A. Rich. 1988. J.Biol.Chem. **263**: 17872-17879.

Fried, M.G.1989. Electrophoresis **10**: 366-376.

Fuller-Pace, F.V., L.R. Bullas, H. Delius, and N.E. Murray. 1984. Proc.Natl.Acad.Sci.USA. **81**: 6095-6099.

Fuller-Pace, F.V., G.M. Cowan, and N.E. Murray. 1985. J.Mol.Biol. **186**: 65-75.

Fuller-Pace, F.V., and N.E. Murray. 1986. Proc.Natl.Acad.Sci.USA. **83**: 9368-9372.

Furst, J.P., W. Pansegrau, R. Frank, H. Blocker, P. Scholz, M. Bagdasarain, and E.Lanke. 1986. Gene **48**: 119-131.

Gann, A.A.F., A.J.B. Campbell, J.F. Collins, A.F.W. Coulson, and N.E. Murray. 1987. Mol.Microbiol. **1**: 13-22.

Glover, S.W., J.Schell, N. Symonds, and K.A. Stacey. 1963. Genet. Res. **4**: 480-482.

Glover, S.W. 1970. Genet.Res.Camb. **15**: 237-250.

Gorbalenya A.E., and E.V. Koonin. 1991. FEBS Letts. **291**: 277-281.

Gough, J.A., and N.E. Murray. 1983. J.Mol.Biol. **166**: 1-19.

Grossweiner, L.I., A.G. Kaluskar, J.F. Baugher. 1976. Int.J.Radiat.Biol.Relat. Stud.Phys.Chem.Med. **29**: 1-16.

Gubler, T., and T.A. Bickle. 1991. EMBO J. **10**: 951-957.

Gubler, M., D. Braguglia, J. Meyer, A. Piekarozicz, and T.A. Bickle. 1992. EMBO J. **11**: 233-240.

Guschlbauer, W. 1988. Gene **74**: 211-214.

Guyot, J.B., J. Grassi, U. Hahn, W. Guschlbauer. 1993. Nucl.Acid.Res. **21**: 3183-3190.

Haberman, A. 1974. J.Mol.Biol. **89**: 545-563.

Hadi, S.M., T.A. Bickle, and R. Yuan. 1975. J.Biol.Chem. **250**: 4159-4164.

Hardy, L.W., J.S. Finer-Moore, W.R. Montfort, M.O. Jo, D.V. Santi, and R.M. Stroud. 1987. Science **232**: 448-455.

Hattman, S., J. Wilkinson, D. Swinton, D. Schlagman, P.M. MacDonald, and G. Mosig. 1985. J.Bacteriol. **164**: 932-937.

Hendrickson, W., and R.F. Schleif. 1984. J.Mol.Biol. **174**: 611-628.

Horiuchi, K., and N.D. Zinder. 1972. Proc.Natl.Acad.Sci.USA. **69**: 3220-3224.

Hubacek, J. 1973. J.Gen.Microbiol. **79**: 257-264.

Ives, C.L., P.D. Nathan, and, J.E. Brooks. 1992. J.Bacteriol. **174**: 7194-7201.

Ingrosso, D., A.V. Fowler, J. Bleibaum, and S. Clarke. 1989. J.Biol.Chem. **264**: 20131-20139.

Kan, N.C., J.A. Lautenberger, M.E. Edgell, and C.A. Hutchison III. 1979. J.Mol.Biol. **130**: 191-209.

Kannan, P., G.M. Cowan, A.S. Daniel, A.A.F. Gann, and N.E. Murray. 1989. J.Mol.Biol. **209**: 335-344.

Kauc, L., and A. Piekarowicz. 1978. Eur.J.Biochem. **92**: 417-426.

Kelleher, J.E., A.S. Daniel, and N.E. Murray. 1991. J.Mol.Biol. **221**: 431-440.

Kelleher, J.E. 1990. Ph.D Thesis. Dept. of Molecular Biology, Edinburgh University, U.K.

Klimasaukas, S.,A. Timinskas, S. Menkevicius, D. Butkiene, V. Butkus, and A. Janulaitis. 1989. Nucl.Acid.Res. **17**: 9823-9832.

Klimasaukas, S., J.L. Nelson, and R.J. Roberts. 1991. Nucl.Acid.Res. **19**: 6183-6190.

Klimasaukas, S., S, Kumar, R.J. Roberts, and X. Cheng. 1994. Cell **76**: 357-369.

Kossykh, V.G., S.L. Schlagman, S. Hattman. 1993. Nucl.Acid.Res. **21**: 3563-3566.

Kossykh, V.G., S.L. Schlagman, S. Hattman. 1994. In **The Third New England Biolabs Workshop on Biological DNA Methylation**. Vilnius, Lithuania.

Kühnlein, U., and W. Arber. 1972. J.Mol.Biol. **63**: 9-19.

Labahn, J., J. Granzin, G. Schluckebier, D.P. Robinson, W.E. Jack, I. Schildkraut, and W. Sanger. 1994. Proc.Natl.Acad.Sci.USA, In Press.

Lakowicz, J.R. 1983. In **Principles of Fluorescence Spectroscopy**. Plenum Press, New York.

Laemmli, U.K. 1970. Nature **227**: 680-685.

Lauster, R. A. Kriebardis, and W. Guschlbauer. 1987. FEBS Lett. **220**: 167-176.

Lauster, R. 1989. J.Mol.Biol. **206**: 313-321.

Lautenberger, J.A., and S. Linn. 1972. J.Biol.Chem. **247**: 6176-6182.

Lautenberger, J.A., N.C. Kan, D. Lackey, S. Linn, M.H. Edgell, and C.A. Hutchinson III. 1978. Proc.Natl.Acad.Sci.USA. **75**: 2271-2275.

Leatherbarrow, R.J. 1990. Grafit version 2.0, Erithacus software Ltd., Staines, UK.

Lederberg, S. 1957. Virology **3**: 496-513.

Linn, S., and W. Arber. 1968. Proc.Natl.Acad.Sci.USA. **59**: 1300-1306.

Loenen, W.A.M., and N.E. Murray. 1986. J.Mol.Biol. **190**: 11-22.

Loenen, W.A.M., A.S. Daniel, H.D. Braymer, and N.E. Murray. 1987. J.Mol.Biol. **198**: 159-170.

Maniatis, T., A. Jeffrey, and D.G. Kleid. 1975. Proc.Natl.Acad.Sci.USA. **72**: 1184-1188.

Meisel, A., T.A. Bickle, D.H. Kruger, and C. Schroeder. 1992. Nature **355**: 467-469.

Meister, J., M. MacWilliams, P. Hübner, H. Jütte, E. Skrzypek, A. Piekarowicz, and T.A. Bickle. 1993. EMBO J. **12**: 4585-4591.

Meselson, M., and R. Yuan. 1968. Nature **217**: 1110-1114.

Mi, S., and R.J. Roberts. 1992. Nucl.Acid.Res. **20**: 4811-4816.

Mi, S., and R.J. Roberts. 1993. Nucl.Acid.Res. **21**: 2459-2464.

Mitraki, M., and L. King. 1992. FEBS Lett. **20**: 307-312.

Moore, M.H., J.M. Gulbis, E.J. Dodson, B. Demple, and P.C.E. Moody. 1994. EMBO J. **13**: 1495-1501.

Murray, N.E., P.L. Batten, and K. Murray. 1973. J.Mol.Biol. **81**: 395-407.

Murray, N.E., J.A. Gough, B. Suri, and T.A. Bickle. 1982. EMBO J. **1**: 535-539.

Murray, N.E., A.S. Daniel, G.M. Cowan, and P.M. Sharp. 1993. Mol.Microbiol. **9**: 133-143.

Nagaraja, V., M. Stieger, C. Nager, S.M. Hadi, and T.A. Bickle. 1985a. Nucl.Acid.Res. **13**: 389-399.

Nagaraja, V., J.C.W. Shepherd, T. Pripfl, and T.A. Bickle. 1985b. J.Mol.Biol. **182**: 579-587.

Nagaraja, V., J.C.W. Shepherd, and T.A. Bickle. 1985c. Nature **316**: 371-372.

Pace, C.N. 1986. Methods Enzymol. **131**: 266-280.

Pace, C.N., B.A. Shirley, J.A. Thomson. 1989. In **Protein Structure: A Practical Approach** (T.E. Creighton, ed.). IRL Press, Eynsham, England. pp. 311-330.

Pogolotti, A.L., A. Ono, R. Subramaniam, and D.V. Santi. 1988. J.Biol.Chem. **263**: 7461-7464.

Posfai, J., A.S. Bhagwat, A.S. Posfai, and R.J. Roberts. 1989. Nucl.Acid.Res. **17**: 2421-2435.

Powell, L.M., D.T.F. Dryden, D.F. Willcock, R.H. Pain, and N.E. Murray. 1993. J.Mol.Biol. **234**: 60-71.

Prakash-Cheng, A. and J. Ryu. 1993. J.Bacteriol. **175**: 4905-4906.

Prakash-Cheng, A., S.S.Chung, and J.Ryu. 1993. Mol.Gen. Genet. **241**: 491-496.

Price, C., and T.A. Bickle. 1986. Microbiol.Sci. **3**: 296-299.

Price, C., T.Pripfl, and T.A. Bickle. 1987. Eur.J.Biochem. **167**: 111-115.

Rafferty, J.B., W.S. Somers, I. Saint-Girons, and S.E.V. Philips. 1989. Nature **341**: 705-710.

Ravetch, J.V., K. Horiuchi, and N.D. Zinder. 1978. Proc.Natl.Acad.Sci.USA. **75**: 2266-2270.

Reich, N.O., and E.A. Everett. 1990. J.Biol.Chem. **265**: 8929-8934.

Reich, N.O., and N. Mashoon. 1991. Biochemistry **30**: 2933-2939.

Rigby, P.W.J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. J.Mol.Biol. **113**: 237-251.

Rossmann, M.G., D. Moras, K.W. Olsen. 1974. Nature **250**: 194-199.

Rubin, R.A., and P. Modrich. 1977. J.Biol.Chem. **252**: 7265-7272.

Rydberg, B., N. Spurr, and P. Karran. 1990. J.Biol.Chem. **265**: 9563-9569.

Sain, B., and N.E. Murray. 1980. Mol.Gen.Genet. **180**: 35-46.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. Proc.Natl.Acad.Sci.USA. **74**: 5463-5467.

Sanger, F., A.R. Coulson, B.G. Barrel, A.J.H. Smith, and B.A. Rao. 1980. J.Mol.Biol. **143**: 161-178.

Santi, D.V., A. Norment, and C.E. Garret. 1984. Proc.Natl.Acad.Sci.USA. **81**: 6933-6977.

Shi, L., D.R. Palleros, A.L. Fink. 1994. Biochemistry **33**: 7536-7546.

Slavik, J. 1982. Biochim. Biophys. Acta. **694**: 1-25.

Smith, H.O., and K.W. Wilcox. 1970. J.Mol.Biol. **51**: 379-391.

Sober, H.A., 1970.In **Handbook of Biochemistry**. 2nd Ed., CRC Press Inc., Boca Raton, USA.

Som, S., and S. Friedman. 1990. J.Biol.Chem. **265**: 4278-4283.

Somers, W.S., and S.E.V. Philips. 1992. Nature **359**: 387-393.

Sommer, R., and H. Schaller. 1979. Mol.Gen.Genet. **168**: 331-335.

Studier, F.W., and P.K. Bandyopadhyay. 1988. Proc.Natl.Acad.Sci.USA. **85**: 4677-4681.

Subbaramaiah, K., and S.A. Simms. 1992. J.Biol.Chem. **267**: 8636-8642.

Sugisaki, H.,K. Kita, and M. Takanami, 1989. J.Biol.Chem. **264**: 5757-5761.

Suri, B., V. Nagaraja, and T.A. Bickle. 1984a. Curr.Top.Microbiol.Immunol. **108**: 1-9.

Suri, B., J.C.W. Shepherd, and T.A. Bickle. 1984b. EMBO J. **3**:575-579.

Suri, B., and T.A. Bickle. 1985. J.Mol.Biol. **186**; 77-85.

Takano, T., T. Watanabe, and T. Fukasawa. 1966. Biochem.Biophys.Res.Comm. **25**: 192-198.

Tao, T., J.C. Bourne, and R.M. Blumenthal. 1991. J.Bacteriol. **173**: 1367-1375.

Taylor, I., J. Patel, K. Firman, and G. Kneale. 1992. Nuc.Acids Res. **20**: 179-186.

Taylor, I., D. Watts, and G. Kneale. 1993. Nucl.Acid.Res. **21**: 4929-4935.

Thatcher, D.R., A. Hitchcock. 1994. In **Mechanisms of Protein Folding** (R.H. Pain, ed.). IRL Press, Eynsham, England. pp. 232-235.

Tyndall, C., J. Meister, and T.A. Bickle. 1994. J.Mol.Biol. **237**: 266-274.

Verdine G. L. 1994. Cell **76**: 197-200.

Vershon, A.K., P. Youderian, M.A. Weiss, M.M. Susskind, and R.T. Sauer. 1985. In **Specificity in Transcription and Translation** (R. Calendar and L. Gold, eds.). Alan Liss, Inc. USA. pp. 209-218.

Vidgren, J., L. Anders-Svensson, and A. Liljas. 1994. Nature **368**: 354-358.

von-Wronski, L. 1991. J.Biol.Chem. **266**: 1064-1070.

Vrielink, A., R. Wolfgang, H.P.C. Driessen, and P.S. Freemont. 1994. EMBO J. **13**: 3413-3422.

Walker, J.E., M. Saraste, M.J. Runswick, and N.J. Gay. 1982. EMBO J. **1**: 945-951.

Weiserova, M., P. Janscak, O. Benada, J. Hubacek, V.E. Zinkevich, S. Glover, and K. Firman. 1993. Nucl.Acids.Res. **21**: 373-379.

Wenzel, C., M. Moulard, A. Lobner-Olesen, and W. Guschlbauer. 1991. FEBS Lett. **280**: 147-151.

Wilke, K., E. Rauhut, N. Noyer-Weidner, R. Lauster, P. Pawlek, B. Behrens, and T.A. Trautner. 1988. EMBO J. **7**: 2601-2609.

Willcock, D.F., D.T.F. Dryden, and N.E. Murray. 1994. EMBO J. **13**: 3902-3908.

Wilson, G.G., 1991. Nucl.Acids.Res. **19**: 2539-2566.

Wilson, G.G., and N.E. Murray. 1991. Ann.Rev.Genet. **25**: 585-627.

Wood, W.B. 1966. J.Mol.Biol. **16**: 118-133.

Wu, J.C., and D.V. Santi. 1987. J.Biol.Chem. **262**: 4778-4786.

Wyszynski, M.W., S. Gabbara, and A.S. Bhagwat. 1991. Nucl.Acids.Res. **20**: 319-326.

Yanish-Perron, C., J. Viera, and J. Messing. 1985. Gene **33**: 103-119.

Yu, P.H., 1983. Biochim Biophys Acta. **742**: 517-524.

Yuan, R., T.A. Bickle, W. Ebbers, and C. Brack. 1975. Nature **256**: 556-560.

Yuan, R., D.L. Hamilton, and J. Burckhardt. 1980. Cell **20**: 237-244.

Zabeau, M., S. Friedman, M. Van Montagu, and J. Schell. 1980. Mol.Gen.Genet. **179**: 63-73.

Zinkevich, V.E., M. Weiserova, V.M. Kryukov, and J. Hubacek. 1990. Gene **90**: 125-128.

Zinkevich, V., P. Heslop, S.W. Glover, M. Weiserova, J. Hubacek, and K. Firman. 1992. J.Mol.Biol. **227**: 597-601.

Zoller, M.J., and M. Smith. 1983. Methods Enzymol. **100**: 468-500.